

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2005, 19:09:59 ; Search time 10473 Seconds
(without alignments)
13400.455 Million cell updates/sec

Title: US-10-614-524-1
Perfect score: 3687
Sequence: 1 ttgacttcaaataggaaaaa.....tactccttatggagggaatag 3687

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_hc.*
- 4: gb_est3.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_gss1.*
- 9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	1.7	880	8	AZ669474
2	59	1.6	908	8	AZ548467
3	57.4	1.6	843	8	AZ551618
4	55.8	1.5	748	9	BX137558
5	55.2	1.5	931	8	BH160272
6	52.8	1.4	1101	9	CNS0039G
7	52.4	1.4	877	8	AZ531291
8	52	1.4	849	8	AZ546009
9	51.8	1.4	886	5	BX422107
10	51.8	1.4	1101	9	CNS00EVL
11	51.2	1.4	900	8	AZ549980
12	51	1.4	1101	9	CNS00EMH
13	51	1.4	1337	9	AG280055
14	50.6	1.4	912	8	AZ551092
15	50.2	1.4	467	4	BM163521
16	49.8	1.4	1253	9	AG289789
17	48.8	1.3	1064	4	BM416119
18	48.6	1.3	890	8	AZ530768
19	48.4	1.3	543	4	BI397189
20	48.4	1.3	905	8	AZ550256
21	48.4	1.3	1101	9	CNS0106X
22	47.6	1.3	828	8	AZ669217
23	47.6	1.3	906	8	AZ529180
24	47.6	1.3	939	8	AZ540454

25	47.6	1.3	1101	9	CNS0100X
26	47.4	1.3	654	9	BM164708
27	47.4	1.3	997	9	CNS005TE
c 28	47	1.3	811	4	BJ403374
c 29	47	1.3	976	8	BH149983
30	46.8	1.3	531	4	BJ363051
31	46.8	1.3	568	4	BJ363776
32	46.8	1.3	582	4	BJ366397
c 33	46.8	1.3	782	9	AG444032
c 34	46.8	1.3	1101	9	CNS0026Z
c 35	46.8	1.3	1307	9	AG347197
36	46.6	1.3	537	5	BU498625
37	46.6	1.3	548	4	BI670654
38	46.6	1.3	556	4	BI814240
39	46.6	1.3	585	5	BU496406
c 40	46.6	1.3	921	8	BH149893
c 41	46.4	1.3	680	4	BJ384484
c 42	46.4	1.3	683	4	BJ386441
c 43	46.4	1.3	826	4	BJ383101
c 44	46.4	1.3	875	8	AZ671853
c 45	46.2	1.3	609	4	BJ444006

ALIGNMENTS

RESULT 1
AZ669474
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AZ669474
ENTY88TR Entamoeba histolytica Sheared DNA
genomic, genomic survey sequence.
A2669474
GSS.
Entamoeba histolytica
Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
1 (bases 1 to 880)
Loftus, B., Van Aken, S. and Fraser, C.
Determination of clone end sequences from Entamoeba histolytica
HM1:IMSS sheared DNA library
Unpublished (2000)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjlftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
DNA library
Seq primer: M13-Reverse
Class: shotgun
High quality sequence start: 22
High quality sequence stop: 853.
Location/Qualifiers
1..880
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/notes="Vector: pHO51; Site 1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, I.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome

Sequencing: A Practical Approach, eds. M. Vaudin and B. Barell, Oxford University Press, 1999).

ORIGIN	Query Match	1.7%	Score 61;	DB 8;	Length 880;
	Best Local Similarity	46.4%;	Pred. No. 0.00038;		
	Matches 199;	Conservative 0;	Mismatches 230;	Indels 0;	Gaps 0;
QY	3252	GTACAAAGAGGGATATGAGACAGGGTTGTGTAAACGATCCATGAAATCGAGAACCAATACAGA	3311		
Db	12	GTAAAGGAGAAGAAGATGAAGATGATGATGAAGAAGAAGATGCAAGATGATGATGAAGAAGA	71		
QY	3312	CGAACTAAATTTAAAAACTGTGAAGAAGAGGAGTGTATCCACCGATACACGAACGGTG	3371		
Db	72	AGATGATGATGAAGAAGAAGATGAAGAAGAAGAAGATGATGAAGAAGATGATGAAGAAGA	131		
QY	3372	TAATGATTATACTGCACACCAAGGTACAGCAGTATGTAAATCCCGTAATCCTCGATATGA	3431		
Db	132	AGAAGATGATGATGAAGACGACGAATATGAATTAGAAGATGATGAAGAAGATGATGAAGA	191		
QY	3432	GGATGCATATGAAGTTTGATACTACAGCATCTGTTAATTACAAACCGACTTTATGAAGAAGA	3491		
Db	192	AGATGATGAAGAAGAAGATGATGAAGAAGAAGAAGATGATGATGATGATGATGAAGAAGA	251		
QY	3492	AACGTATACAGATGTACGAAGAGATATCATTTGTGAATATGACAGAGGGTATGTGAATTA	3551		
Db	252	TGATGATGAAGATGATGAAGAAGATGAAGATGATGAAGAAGATGAAGATGATGAAGAATGA	311		
QY	3552	TCCACCACTACCGCTGGTTATATGACAAAGAATTAGAATACTTCCCGAAACCGATAA	3611		
Db	312	TGAAGATGATGAAGATGATGAAGATGATGAAGACGACGAATATGAATTAGAGATGAAGA	371		
QY	3612	GGTATGATTGAGATTGGAGAAAACGGAAGGGAAGTTTATTGTAGACAGCGCTGGAAATTACT	3671		
Db	372	TGATGAAGATGATCAAGATGATGATGAAGACGAATATGAATTAGAGATGATGATGATTATT	431		
QY	3672	CCTTATGGA	3680		
Db	432	CTATGATGA	440		

RESULT 2	908 bp	linear	GSS 14-NOV-2000
AZ548467			
LOCUS	908 bp	DNA	
DEFINITION	Entamoeba histolytica	Sheared DNA	Entamoeba histolytica
	ENTEX30TR		
	genomic		
	genomic survey sequence.		

A2548467.1 GI:11172102
 Entamoeba histolytica
 Entamoeba histolytica
 Eukaryota; Entamoebidae; Entamoeba.
 1 (bases 1 to 908)
 Loftus, B., Van Aken, S. and Fraser, C.
 Determination of clone end sequences from Entamoeba histolytica
 HM1:IMSS sheared DNA library
 Unpublished (2000)
 Contact: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 912 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: bjoftus@tigr.org
 Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
 DNA library
 Seq primer: M13-Reverse
 Class: shotgun
 High quality sequence start: 17
 High quality sequence stop: 828.
 Location/Qualifiers
 1..908
 /organism="Entamoeba histolytica"
 source

```

/mol_type="genomic DNA"
/strain="HM1: IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOS1; Site 1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. parasitol.
77-450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaubin and B.
Barrell Oxford University Press, 1999)."

```

ORIGIN

Query Match	1.6%	Score 59	DB 8	Length 908
Best Local Similarity	44.2%	Pred. No. 0.0013		
Matches 337	Conservative 0	Mismatches 420	Indels 6	Gaps 2
Qy	2793	ACGAGGAAAAATGAGAGACAAACGTGAAACCTACAAATTCGAAACAAACGAGTATA	2852	
Db	66	AGAAGATGATGATGAAGAAGATGATGATGAAGAAGATGATGAAGAAGATGATGAAGA	125	
Qy	2853	TACAGAGGCAAAAGAGCTGTGATGCTTTATTTGTGTAGATTTCTCAATATAATAGATTACA	2912	
Db	126	AGAAGAAGATGATGAAGATGAAGATGATGAAGACGATGAAGACGACGAATATGAATTAGA	185	
Qy	2913	AGCGGATACAAACATTTGGCATGATTCATGGCGAGATAAACTTTGTCATCGAATTCGAGA	2972	
Db	186	AGATGATG---ATCATGAAGAAGAAGATGATGATGAAGATGATGAAGACGATGAAGA	242	
Qy	2973	GGCTTATCTGCAGAAATATCTCTTATCCCGGTGTAAATCGGAAATTTTTCGAAGAATT	3032	
Db	243	CGACGAATATGAATTTAGAAGATGATGATGATGAAGAAGAAGATGATGATGAAGAAGA	302	
Qy	3033	AGAAGGTGCGATTATCACTGCAATCTCCCTATACGATCGAGAAATGCTCTTAAAAATGG	3092	
Db	303	AGAAGATGATGAAGATGATGATGATGAAGAGAAGATGAAGAAGATGATGAAGAAGATGA	362	
Qy	3093	TGATTTTAAATGGAATAGCATGCTGGAAATGTAAAGGGGCATGTAGATGTTACAAACAGAG	3152	
Db	363	TGAAGATGATGATGAAGAAGAAGATGATGATGAAGATGATGAAGATGATGAAGACGA	422	
Qy	3153	CCATCACCGTTCTGCTCTTATCCAGAAATGGGAAGCAGAGTGTCAACAGCAGTTCG	3212	
Db	423	CGAATATGAATTTAGAAGATGATGATGAAGAAGAAGAAGATGATGATGAAGAAGAAGA	482	
Qy	3213	CGTCTGTCGGGGCGTGGCTATATCTCTCTGTCAACGCTACAAGAGGGATATGGAGA	3272	
Db	483	TGATGATGAAGATGATGATGAAGAAGAAGATGAAGATGATGATGAAGAAGAAGATGAAGA	542	
Qy	3273	GGGTGTGTAAACGATCCATGAAATCGAGACA---ATACAGACGAACTAAAAATTTAAAAA	3329	
Db	543	TGATGATGAAGAAGAAGATGATGATGAAGAAGAAGATGAAGAAGAAGAAGATGATGAAGA	602	
Qy	3330	CTGTGAAGAAGAGGAAGTGTATCCAACGGATACAGGAACCTGTAAATGATTTACTGCACA	3389	
Db	603	AGATGATGAAGAAGAAGAAGATGATGATGAAGACGACGAATATGAATTTAGAAGATGATGA	662	
Qy	3390	CCAGGTACAGCAGTATGTAAATTTCCCGTAATCTGGATATGAGGATGCATATGAAGTTGA	3449	
Db	663	AGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGAAGAAGATGATGATGA	722	
Qy	3450	TACTACAGCATCTGTTAAATTCAAAACCGACTTATGAAGAAGAAACGTTATACAGATGTACG	3509	
Db	723	TGATGATGAAGAAGATGATGATGAAGATGATGAAGAAGATGAAGAAGATGATGAAGATGA	782	
Qy	3510	AAGAGATAATCAATGTGAATATGACAGAGGGTATGTGAATTAT	3552	

```

Db      783 AGAAGATGAAGATGATGAAGATGATGAAGATGATGAAGATGAT 825

RESULT 3
AZ551618/c
LOCUS   AZ551618      843 bp      DNA      linear      GSS 14-NOV-2000
DEFINITION   ENT54TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
              genomic, genomic survey sequence.
ACCESSION   AZ551618
VERSION     1
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica
ORGANISM    Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1 (bases 1 to 843)
AUTHORS    Loftus,B., Van Aken,S. and Fraser,C.
TITLE      Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
JOURNAL    Unpublished (2000)
COMMENT    Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjoftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Reverse
            Class: shotgun
            High quality sequence start: 39
            High quality sequence stop: 838.
            Location/Qualifiers
                1..843
                    /organism="Entamoeba histolytica"
                    /mol_type="genomic DNA"
                    /strain="HM1:IMSS"
                    /db_xref="taxon:5759"
                    /clone_lib="Entamoeba histolytica Sheared DNA"
                    /note="Vector: PHOS1; Site 1: Bst I; Constructed at The
                    Institute for Genomic Research (TIGR), Rockville, MD.
                    Genomic DNA isolated from broth cultures of E. histolytica
                    using a method described by Clark and Diamond (Clark,
                    C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
                    method for isolate identification. Exp. Parasitol.
                    77:450.). The DNA was mechanically sheared to give a
                    tight size distribution (~2 kb). The v + i method used for
                    the library construction is described in detail in Smith,
                    H.O. and Venter, J.C. (Making small insert libraries for
                    whole genome shotgun sequencing projects. In Genome
                    Sequencing: A Practical Approach, eds. M. Vaudin and B.
                    Batell, Oxford University Press, 1999).".

FEATURES             source
    source
    1..748
        /organism="Danio rerio"
        /mol_type="genomic DNA"
        /db_xref="taxon:7955"
        /clone="DKEY-95C15"
        /tissue_type="Testis"
        /note="vector pindigobAC-536"

ORIGIN
Query Match      1.6%; Score 57.4; DB 8; Length 843;
Best Local Similarity 43.8%; Pred. No. 0.0032;
Matches 348; Conservative 0; Mismatches 441; Indels 6; Gaps 2;

QY 2793 AGCAGAGAAAATGGAGACAAACCTGAAACACTCAATTCGAACAAACGAGTATA 2852
DB 812 AGAAGATGATGATGAAGAAGATGATGATGAAGAAGATGATGAAGAAGATGATGAAGA 753
QY 2853 TACAGAGCGCAAAAGAGCTGGAGTCTTTATTTGTAGATTCTCAATATAATAGATTACA 2912
DB 752 AGAAGATGATGATGAAGAAGATGATGATGAAGAAGATGATGAAGAAGATGATGAAGA 693
QY 2913 AGCGGATACAAACATTTGGCATGATTCATCGGCGCAGATAAACTTGTTCATCGAATTCGAGA 2972
DB 692 AGATGAT--GATGATGAAGAAGATGATGATGATGAAGAAGATGATGAAGAAGATGATGAAGA 636
QY 2973 GCCTTATCTGTGAGAAATTAATCTGTTATCCCGGGTGTAAATCCGGGAAATTTTGAAGATT 3032
DB 635 CGACGAAATATGAAATTAAGATGATGATGATGAAGAAGATGATGATGAAGAAGATGATGAAGA 576

```

```

QY 3033 AGAAGTTCGCATTATCACTGCAATCTCCCTATACGATCGCAGAAATGTCTGTTAAAAATGG 3092
DB 575 AGAAGATGATGATCAAGATGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGA 516
QY 3093 TGATTTTAAATAGGATTAGCATCTGGAATGTAAAGAGGCATCTAGATGTACAAACAGAG 3152
DB 515 TGAAGATGATGATGAAGAAGATGATGATGAAGAAGATGATGATGAAGAAGATGATGAAGA 456
QY 3153 CCATCACCGTTCCTGCTCTTATCCAGAAATGGGAAGCAGAAAGTGTCAACAAGCAGTTCCG 3212
DB 455 CGAATATGAATTAGAAGATGATGATGAAGAAGATGATGATGAAGAAGATGATGAAGAAGAA 396
QY 3213 CGTCTGCCGGGGCGTGTATATCTCCCTGTTCACACCGGTACAAAGAGGATATGAGAGA 3272
DB 395 TGATGATGAAGATGATGATGAAGAAGATGATGATGAAGAAGATGATGATGAAGAAGATGA 336
QY 3273 GGGTGTGTAAACGATCCATCAATTCGAGA--ACAATACAGACGAACCTAAATTTTAAAAA 3329
DB 335 TGATGATGAAGAAGATGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGA 276
QY 3330 CTGTGAAGAAGAGAGTGTATCCACGATACAGGAACCGTGAATGATTATATCTACTGCACA 3389
DB 275 AGATGATGAAGAAGATGATGATGAAGAAGATGATGATGAAGAAGATGATGAAGAAGATGA 216
QY 3390 CCAAGGTACAGCAGTATGTAATTCCTCCGTATGCTGGATATGAGGATGCATATGAAGTTGA 3449
DB 215 AGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGA 156
QY 3450 TACTACAGCATCTTTAATTACAAACCGACTTATGAAGAGNAACGTTACAGATGTACG 3509
DB 155 TGATGATGAAGAAGATGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGA 96
QY 3510 AAGAGATAATCATTGTGAATATCACAGCGGTATGTGAATTTATCCACCTACCAGCTGG 3569
DB 95 AGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGATGAAGAAGATGAATATA 36
QY 3570 TTATATGACAAAAGA 3584
DB 35 TTAGAAGATGAATA 21

```

```

RESULT 4
BX137558      748 bp      DNA      linear      GSS 28-JAN-2003
LOCUS   BX137558      748 bp      DNA      linear      GSS 28-JAN-2003
DEFINITION   Danio rerio genomic clone DKEY-95C15, genomic survey sequence.
ACCESSION   BX137558
VERSION     BX137558.1 GI:27968872
KEYWORDS    GSS.
SOURCE      Danio rerio (zebrafish)
ORGANISM    Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1 (bases 1 to 748)
AUTHORS    Humphray,S.J., Huckle,E. and Durham,J.I.
TITLE      Direct Submission
JOURNAL    Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome
            Campus, Hinxton, Cambridgeshire CB10 1SA, UK. E-mail enquiries:
            humquerry@sanger.ac.uk Unpublished
            This sequence was generated from the SP6 end of BAC 95C15. 95C15 is
            part of the Daniokey BAC Library created by R. Plasterk and N.V.
            keygene. Further details:
            http://www.sanger.ac.uk/Projects/D_rerio/.
FEATURES             Location/Qualifiers
    source
    1..748
        /organism="Danio rerio"
        /mol_type="genomic DNA"
        /db_xref="taxon:7955"
        /clone="DKEY-95C15"
        /tissue_type="Testis"
        /note="vector pindigobAC-536"

ORIGIN

```

Query Match 1.5%; Score 55.8; DB 9; Length 748;
 Best Local Similarity 45.8%; Pred. No. 0.008;
 Matches 192; Conservative 0; Mismatches 227; Indels 0; Gaps 0;

QY 3238 CTCGGTGTACAGCGTCAAAAGAGGGGATATGGAGAGGGTTGTGTAAACGATCCATGAAATC 3297
 Db 130 CTCGAAGATTTCAGAGAAGAAAGATTTCAGAGAAGAAAGATTTCAGAGAAGAAAGATTTCAGAGA 189

QY 3298 GAGACATATACAGCAGTAAATTTAAAACTGTGAGAGAGGAGGAGTGTATCCACG 3357
 Db 190 GATTTCAGAGAAGATTTCAGAGAAGATTTCAGAGAAGAAAGATTTCAGAGAAGAAAGAA 249

QY 3358 GATACAGGAACGTGTATATGTTACTGTCACACCAAGGTACAGCAGTATGTAATTCCTCGT 3417
 Db 250 GATTTCAGAGAAGATTTCAGAGAAGATTTCAGAGAAGAAAGATTTCAGAGAAGAAAGATTTCAGAA 309

QY 3418 AATGCTGGATATGAGGATGATATGAAGTTGATATCTACAGCATCTGTTAATTCACAAACG 3477
 Db 310 GAAGATTTCAGAGAAGATTTCAGAGAAGAAAGATTTCAGAGAAGAAAGATTTCAGAGAAGATTTC 369

QY 3478 ACTTATGAGAGAAGAAAGTATACAGATGTACGAGAGAGATATCATTTGTGATATGACAGA 3537
 Db 370 GAAGATTTCAGAGAAGATTTCAGAGAAGAAAGATTTCAGAGAAGAAAGATTTCAGAGAAGATTTC 429

QY 3538 GGGTATGTGATATTCACCACTACAGCTGGTTTATATGACAAAGAAATTAGAATCTTC 3597
 Db 430 GAAGAAGAAAGATTTCAGAGAAGATTTCAGAGAAGATTTCAGAGAAGAAAGATTTCAGAGAAGAA 489

QY 3598 CCAGAAACCGATAGGATGATGATGAGATTGGAGAAACGGAAGGAAGTTTATGTTAGA 3656
 Db 490 GAAGATTTCAGAGAAGATTTCAGAGAAGATTTCAGAAATAGATTTCAGAGAAGATTTCAGAGA 548

RESULT 5
 BH160272 931 bp DNA linear GSS 24-SEP-2001
 LOCUS ENT0V49TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
 DEFINITION genomic, genomic survey sequence.

ACCESSION BH160272.1 GI:157333710
 VERSION GSS.
 KEYWORDS Entamoeba histolytica
 SOURCE Entamoeba histolytica
 ORGANISM Entamoeba histolytica
 Eukaryota; Entamoebidae; Entamoeba.

REFERENCE 1 (bases 1 to 931)
 Loftus, B., Wang, Z., Van Aken, S. and Fraser, C.
 Determination of clone end sequences from Entamoeba histolytica
 HMI:IMSS sheared DNA library (2001)
 Unpublished (2001)

JOURNAL
 CONTACT: Brendan J Loftus
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: bjl@fuser.tigr.org
 Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
 DNA library
 Seq primer: M13-Reverse
 Class: shotgun
 High quality sequence start: 8
 High quality sequence stop: 677.
 Location/Qualifiers

FEATURES
 source 1..931
 /organism="Entamoeba histolytica"
 /mol_type="genomic DNA"
 /strain="HMI:IMSS"
 /db_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /note="Vector: pHOS1; Site 1: Bst I; Constructed at The
 Institute for Genomic Research (TIGR), Rockville, MD.
 Genomic DNA isolated from broth cultures of E. histolytica

using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Borell, Oxford University Press, 1999)."

ORIGIN

Query Match 1.5%; Score 55.2; DB 8; Length 931;
 Best Local Similarity 47.4%; Pred. No. 0.012;
 Matches 165; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

QY 3257 AAGGCGGATGAGAGGGTGTGTAAACGATCCATGAATCGAGAACTACAGACGAAC 3316
 Db 330 AAGAAGAAGATGAAGATGATGAAGAAGAGATGATGATGAAGAAGAAGATG 389

QY 3317 TAAATTTAAAACTGTGAAGAAGAGGAAGTGTATCCAAACGGATACAGGAACGTGTAATG 3376
 Db 390 ATGATGAAGAGAGATGAAGAAGAGAGATGATGAAGAAGAGATGAAGAAGAAG 449

QY 3377 ATPATCTGCACACCAAGGTACAGCAGTATGTAATTCCTGTAATCTGGATATGAGATG 3436
 Db 450 ATGATGATGAAGACGACGAATATGTAATGAAGATGATGAAGAAGATGATGAAGAAGATG 509

QY 3437 CATATCAAGTTGATATACAGCATCTGTTAATTCACAAACCGCTTATGAAGAAGAAACGT 3496
 Db 510 ATGAAGAAGAAGATGATGAAGAAGAAAGATGATGATGATGATGATGAAGAAGATG 569

QY 3497 ATACAGATGTACCAAGAGATTAATCTGTGAATATACAGAGGGTATGTGAATTTATCCAC 3556
 Db 570 ATCAAGATGATGAAGAAGATGAAGATGATGAAGAAGATGAAGATGAAGATGAAG 629

QY 3557 CACTACAGCTGTTATATGACAAAGAATTAGAACTCTCCAGAAA 3604
 Db 630 ATGATGAAGATGATGAAGATGATGAAGACGACGATTTATGATTAGAAA 677

RESULT 6

CNS0039G/c 1101 bp DNA linear GSS 03-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
 DEFINITION BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL063921 GI:4941778
 VERSION AL063921.1
 KEYWORDS GSS.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)
 Genoscope.
 Direct Submission
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

TITLE
 JOURNAL
 COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library

	494	TGATGAAGATGAAGAAGATGATGATGAAGAAGACGAAGAAGATGATGATGATGA	435
Db			
	3108	ATTAGCATGCTGAAATGTAAAAGGGCGATGTAGATGTACAACAGAGCCATCACCGTTTCTGT	3167
Qy			
Db			
	434	ATTTGAATTAGAAGATGAAGAAGATGAAGAAGATGAAGAAGATGATGATGATGATGA	375
Db			
	3168	CCTTGTTATCCCAGAATGGGAAGCAGAGTGTCAAGCAGTTCCGGTCTGTCCGGGGCG	3227
Qy			
Db			
	374	AGATGATGAAGATGATCAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGA	315
Db			
	3228	TGGCTATATCTTCGTGCACACGGTACAAGAGGGCATTGGAGAGGGTGTGTGAACGAT	3287
Qy			
Db			
	314	TGATGATGATGAAGATGAAGAAGATGAAGATGATGATGAATTTGAATTAGAAGATGA	255
Db			
	3288	CCATGAAATCGAACAATAACACACCAACTTAAAAATTTAAAAATCTGCAAGAAGAGGAAGT	3347
Qy			
Db			
	254	AGAGATGAGATGATGATGACAGACCGAAGACGAGAGATGATGATGAATTTGAATTAGA	195
Db			
	3348	GTAATCAAAGGATPACAGGAACGTTGTAATGATTATATCTGCACACCAGGTACACGATATG	3407
Qy			
Db			
	194	AGATGAAGATGAAGATGAAGATGAAGAAGATGAAGATGATGAAGAAGACGAAGAAGA	135
Db			
	3408	TAATTCCTGTAATCTCGTATGAGATGCATATGAAGTTGTACTACAGCATCTGTTAA	3467
Qy			
Db			
	134	AGATGATGATGAAGATGAAGAAGATGATGATGAAGATGATGATGAAGATGATGAAGA	75
Db			
	3468	TTACAAACCCACTTATCAAGAAGA	3491
Qy			
Db			
	74	TGACGAAGATGTTGATGAAGATGA	51
Db			

RESULT 9
BX422107
LOCUS
DEFINITION BX422107 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens CDNA clone CS0DH007YD10 3-PRIME, mRNA sequence.
ACCESSION BX422107
VERSION BX422107.2 GI:46930727
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 886)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 13, 2003 this sequence version replaced gi:30651307.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.

This sequence belongs to sequence cluster 3628.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0DH007DB05NP1&c=3628.r.

FEATURES
source
1..886
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DH007YD10"
/tissue_type="T CELLS (JURKAT CELL LINE)"
/cell_line="JURKAT CELL LINE"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector."

Library was not normalized."

```
ORIGIN
Query Match      1.4%; Score 51.8; DB 5; Length 886;
Best Local Similarity 38.6%; Pred. No. 0.089;
Matches 118; Conservative 34; Mismatches 154; Indels 0; Gaps 0;
QY 1851 AATAAGTATTAAGTAATATGAGGTAGACAAACGTTTTCACCTTTCATTAATTAATTCAT 1910
Db 86 RAAAAAATAGGAARAGGAAGRAAATGAAAMTGTGAANNAANNTTAANNANG 145
QY 1911 TCCAAATCTGCAACCTTCGAAGCAGATACGATTTAGAAAGGCGCAAGAGCGGTGAA 1970
Db 146 GNAATTTAAACNNWTCMMAAAGRAAARAGAAARAGRAAARAGTAATAATATMC 205
QY 1971 TGTCTCTTCTACTAATCGAATCAAGAAAGTTCGAAACGATGTCAGACATTCATAT 2030
Db 206 TGAATATTTGCAAAATATWATCAAGGGAAGAAAGAAATGGAACACAKATMAAA 265
QY 2031 TGATCAAGTATCAATTTAGTGGCGTGTTCATCGGATGAATTCGTTAGATGAAGAG 2090
Db 266 AATDTTADAADAAAKKDATATATWAAAAAAAKATKAKATTTGKAAGAAWAAWAA 325
QY 2091 AGAATTAATCTGAGAAAGTGAATATCGAAGCACTCAGTGATGAAAGAACTTACTCCA 2150
Db 326 TGGATGAWRAAARAAATAGAAWAAACCTTATTATTARAATARAAGARAAARAA 385
QY 2151 AGATCC 2156
Db 386 AAGCC 391

RESULT 10
CNS00EVL      1101 bp DNA linear GSS 04-JUN-1999
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC:
BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1101)
Genoscope.
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammos in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clones="BACR29B23"

FEATURES
source
1..1101
```

/clone_lib="RPCI-98"
/note="end : T7"

ORIGIN

```
Query Match      1.4%; Score 51.8; DB 9; Length 1101;
Best Local Similarity 33.3%; Pred. No. 0.094;
Matches 209; Conservative 318; Mismatches 318; Indels 4; Gaps 2;
QY 1851 AATAAGTATTAAGTAATATGAGGTAGACAAACGTTTTCACCTTTCATTAATTAATTCAT 1910
Db 470 WAAATTTWAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 529
QY 1911 TCCAAATCTGCAACCTTCGAAGCAGATACGATTTAGAAAGGCGCAAGAGCGGTGAA 1970
Db 530 WTTWATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 589
QY 1971 TGTCTCTTCTACTAATCGAATCAAGAAAGTTCGAAACGATGTCAGACATTCATAT 2030
Db 590 WTATAATTTWATTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 649
QY 2031 TGATCAAGTATCAATTTAGTGGCGTGTTCATCGGATGAATTCGTTAGATGAAGAG 2090
Db 650 TTATTAATAATTAATTTWTTWTTAAATTAATAATAATAATAATAATAATAATAATAATA 709
QY 2091 AGAATTAATCTGAGAAAGTGAATATCGAAGCACTCAGTGATGAAAGAACTTACTCCA 2150
Db 710 AAWAATWATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 769
QY 2151 AGATCCAACTTCACATCCAT-CAATAAGCAACGACAGCTTCATATCTACTAATGAGCAAT 2209
Db 770 WTAATWATWAAWWTATATAWATAATAATAATAATAATAATAATAATAATAATAATAATA 829
QY 2210 CGAATTTTACATCTCTCATCAACATCTGAAACATCGATGCTGGGAAGTGAGAACATTA 2269
Db 830 AWAATWATATAWATAATAWAAWAAATTAATAATAATAATAATAATAATAATAATAATAATA 889
QY 2270 CAATCCAGGAAGAAATGACGTATTTAAAGAGAAATTAAGTCACACTACCGGGGACTTTTA 2329
Db 890 WTTTWWAAWATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 949
QY 2330 ATGAGTGTAT--CCGACGTATTTATATCAAAAAATAGGAGAGCTCGGAATTAAGCTT 2386
Db 950 TTAATWTTATWATWATWATWATWATWATWATWATWATWATWATWATWATWATWATWAT 1009
QY 2387 ATATCTCGTACCAATTTAAGAGGGTATATTTGAAGATAGTCAAGATTTAGAGATATATTGA 2446
Db 1010 TATTAATWTTATTTTAAWAAWTAATAATAATAATAATAATAATAATAATAATAATAATA 1069
QY 2447 TCGTTTATTAATGCAACATGAAACAT 2473
Db 1070 TATWTAATAATAWAAWAAWAAWAAWAAWAT 1096

RESULT 11
AZ549980
LOCUS
DEFINITION
ENTD94TF Entamoeba histolytica Sheared DNA linear GSS 14-NOV-2000
genomic, genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Entamoeba histolytica
Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE
1 (bases 1 to 900)
Loftus, B., Van Aken, S. and Fraser, C.
Determination of clone end sequences from Entamoeba histolytica
HM1:IMSS sheared DNA library
Unpublished (2000)
CONTACT:
Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
```

Tel: 301 838 0208
 Fax: 301 838 3543
 Email: bilofus@igrr.org
 Clones are derived from the Entamoeba histolytica HM1:IMSS sheared DNA library
 Seq primer: M13-Forward
 Class: shotgun
 High quality sequence start: 20
 High quality sequence stop: 890.
 Location/Qualifiers
 source
 1. 900
 /organism="Entamoeba histolytica"
 /mol_type="genomic DNA"
 /strain="HM1:IMSS"
 /db_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /notes="vector: pBOS1; Site 1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999)."

FEATURES

source

ORIGIN

Query Match 1.4%; Score 51.2; DB 8; Length 900;
 Best Local Similarity 48.0%; Pred. No. 0.13;
 Matches 146; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

QY 3290 ATGAATCGAGAACAAATACAGCAACTAAATTTAAACTGTGGAAGAGAGAAAGTGT 3349
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 323 ATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGAAG 382
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 3350 ATCAACCGGATACAGGACGTGTAATGATATATCTGCACACCGAGGTACAGCAGTATGTA 3409
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 383 ATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATG 442
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 3410 ATTCCTGTAATCTGGATATGAGATGCATATGAAGTTGATCTACAGCATCTGTTAAAT 3469
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 443 AAGAAGATGAAGATGAAGACGATGATGAAGTATGAAGTATGAAGATGAAGATGAAGATG 502
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 3470 ACAAACCGATTATGAAGAGAAACGTATACAGATGATGAAGAGATGAATCATTTGTAAT 3529
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 503 ATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGAAG 562
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 3530 ATGACAGCGGTATGTAATATCCACTACCGCTGCTTATATGACAAAAGATTAG 3589
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 563 ATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATGATGAAGATG 622
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 3590 AATA 3593
 |||||
 Db 623 AAGA 626
 |||||

RESULT 12

CNS00EMH/c

LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC: BACR29M06 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
 DEFINITION
 accession AL069378.1 GI:4949521
 version AL069378
 keywords GSS.
 source Drosophila melanogaster (fruit fly)
 organism Drosophila melanogaster
 eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1101)

Genoscope.

Direct Submission

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammose in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Location/Qualifiers

1. 1101

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

/clone="BACR29M06"

/clone_lib="RPCI-98"

/notes="end : T7"

ORIGIN

Query Match 1.4%; Score 51; DB 9; Length 1101;
 Best Local Similarity 32.0%; Pred. No. 0.15;
 Matches 102; Conservative 69; Mismatches 148; Indels 0; Gaps 0;

QY 2681 GCGTGTGGGTGTTTCAAGATTAAAGACGAGGAGGTCTATGCAAGACTAGGAAATCTGG 2740
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 932 GCGTGTGAGKAGGATKTKDWTAKAWDTRWTAWAAWATAATATATGATDAATAWTGTGG 873
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 2741 AATTATTAGAGAAACCATTTATTAGGAGAGACACTGCTCTGTGAGAGACAGAGA 2800
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 872 AKTGATGAGDARDRARARAAKGTCTGATGATWKAkakAKAWKKKATGTGKAGATRTD 813
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 2801 AAAAAATGAGAGACAAACGTGAAAAAACAATTCGAAACAAACGAGTATATACAGAGG 2860
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 812 AGAKAGATGTRGAGTATATKATRAARADAGATATAAGGDRATRKARRAKATETAGRA 753
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 2861 CAAAAGAGCTGTGGATGCTTTATTTGTAGATTCTCAATATATAGATTACAAGCGGATA 2920
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 752 KRWAkakAWKAGAGTGWKAWAKATGTGRTATTTKAKWRKARKRWGTAKATDTGRAAAA 693
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 2921 CAAACATTGGCATGATTCATCGGAGAGATAAATCTGTTATCGAATTCGAGAGGCTTATC 2980
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 692 AARTAAAGAGATRTTRRKRWKAWAWATATATAGTGWGTAGAAAAAAGAGAAAGTGTGTG 633
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 2981 TGTCAGAATTATCTGTTAT 2999
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 632 TDGGGKRDRTDGGRWGT 614
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13

AG280055/c

LOCUS

DEFINITION

AG280055

AG280055.1

GI:47852932

KEYWORDS

GSS.

SOURCE

ORGANISM

AG280055 1337 bp DNA linear GSS 02-JUN-2004
 Mus musculus molossinus DNA, clone:MSMg01-051D13.T7, genomic survey sequence.
 accession AG280055
 version AG280055.1
 keywords GSS.
 source Mus musculus molossinus
 organism Mus musculus molossinus
 eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2005, 03:27:14 ; Search time 54 Seconds
(without alignments)
2188.040 Million cell updates/sec

Title: US-10-614-524-2
Perfect score: 6479
Sequence: 1 LTSNRKNEIINALSIPAV.....IGTEGKFVDSVELLMEE 1228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5745	88.7	1228	S00873	parasporal crystal
2	3791	58.5	1176	2 A48970	parasporal crystal
3	3484.5	53.8	1189	2 S00944	parasporal crystal
4	3433.5	53.0	1165	2 S11446	parasporal crystal
5	3433	53.0	1166	2 S32645	parasporal crystal
6	3426.5	52.9	1181	2 A41052	parasporal crystal
7	3421	52.8	1171	2 I40572	parasporal crystal
8	3413.5	52.7	1174	2 S32649	parasporal crystal
9	3413.5	52.7	1176	2 J70241	parasporal crystal
10	3412.5	52.7	1174	2 A42459	parasporal crystal
11	3411	52.6	1160	2 S32647	parasporal crystal
12	3405.5	52.6	1176	2 JC2219	parasporal crystal
13	3400.5	52.5	1176	2 S02215	parasporal crystal
14	3399.5	52.5	1176	2 A22617	parasporal crystal
15	3399.5	52.5	1178	1 USBSKH	parasporal crystal
16	3399	52.5	1171	2 A37829	parasporal crystal
17	3387	52.3	1177	2 A49785	parasporal crystal
18	3323.5	51.3	1172	2 S32689	parasporal crystal
19	3283	50.7	1155	2 JD0002	parasporal crystal
20	3269	50.5	1155	2 A26513	parasporal crystal
21	3268	50.4	1155	2 S02134	parasporal crystal
22	3257.5	50.3	1156	2 A29125	parasporal crystal
23	3251	50.2	1155	2 T39838	parasporal crystal
24	3161.5	48.8	1156	2 A29838	parasporal crystal
25	2434	37.6	934	2 A22798	parasporal crystal
26	2294.5	35.4	1138	2 A48944	parasporal crystal
27	2200.5	34.0	1157	1 S49247	parasporal crystal
28	2170.5	33.5	719	2 I40590	cryV465 protein -
29	2116	32.7	719	2 I39815	insecticidal prote

30	2107	32.5	719	2 S25383	parasporal crystal
31	2101	32.4	719	2 I39814	insecticidal prote
32	2002	30.9	1160	2 I40589	parasporal crystal
33	1947.5	30.1	1154	2 S39536	parasporal crystal
34	1930	29.8	823	2 S04181	parasporal crystal
35	1621	25.0	1156	2 S19306	parasporal crystal
36	1616.5	24.9	1136	1 USBS81	parasporal crystal
37	1511.5	23.3	1180	2 I39870	parasporal crystal
38	1498.5	23.1	1180	2 A26858	parasporal crystal
39	1227	18.9	380	2 B42459	hypothetical prote
40	1164	18.0	655	2 JC7140	pro toxin - Bacillu
41	1064	16.4	652	2 A27323	parasporal crystal
42	1062.5	16.4	659	2 S10228	parasporal crystal
43	1053.5	16.3	652	2 I39811	parasporal crystal
44	999.5	15.4	649	1 JH0261	parasporal crystal
45	987.5	15.2	618	2 S11445	parasporal crystal

ALIGNMENTS

RESULT 1

S00873
parasporal crystal protein cryBa1 - Bacillus thuringiensis subsp. thuringiensis
N;Alternate names: parasporal crystal protein cryA4
C;Species: Bacillus thuringiensis subsp. thuringiensis
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C;Accession: S00873
R;Brizzard, B.L.; Whiteley, H.R.
Nucleic Acids Res. 16, 2723-2724, 1988
A;Title: Nucleotide sequence of an additional crystal protein gene cloned from Bacillus t
A;Reference number: S00873; MUID:88203216; PMID:3362680
A;Accession: S00873
A;Molecule type: DNA
A;Residues: 1-1228 <BRI>
A;Cross-references: UNIPROT:P05517; EMBL:X06711; NID:q40264; PIDN:CAA29898.1; PID:q58094;
C;Genetics:
A;Gene: cryA4
A;Start codon: TTG
C;Superfamily: parasporal crystal protein
C;Keywords: delta-endotoxin

Query Match 88.7%; Score 5745; DB 2; Length 1228;

Best Local Similarity 89.1%; Pred. No. 0;

Matches 1098; Conservative 35; Mismatches 90; Indels 10; Gaps 3;

QY	1	LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVASTVOTGI	60
DB	1	MTSNRKNENIIN-----AVSNHSAQMDLLPDARIEDSLCIAEGNNIDPFVSAIVTQGI	55
QY	61	NIAGRILGLGVFPFAGQIASFYSLVGLWPRGRDOWEIELEHVEQLINQOITENARNTA	120
DB	56	NIAGRILGLGVFPFAGQIASFYSLVGLWPRGRDOWEIELEHVEQLINQOITENARNTA	115
QY	121	LARLQGLGDSFRAYQOSLEWLENRDDARTSRVLVYQYIALELDFLNAMPLFAIRNVEVP	180
DB	116	LARLQGLGDSFRAYQOSLEWLENRDDARTSRVLVYQYIALELDFLNAMPLFAIRNVEVP	175
QY	181	LLMVYQAANLHLLLRDASLFGSEFGLTSQETQRYVERQVETRDYSDYCVENYNTGLN	240
DB	176	LLMVYQAANLHLLLRDASLFGSEFGLTSQETQRYVERQVETRDYSDYCVENYNTGLN	235
QY	241	SLRGTTNAASVRYNQFRDLTLGLVLDLVALFPSTYDTPTINTSAQLTREYVTDATGATG	300
DB	236	SLRGTTNAASVRYNQFRDLTLGLVLDLVALFPSTYDTPTINTSAQLTREYVTDATGATG	295
QY	301	VNMASMNWYNNAPPSAIEATVIRSPHLLDFLEQLTIFSTSSWSATRMTYWRGHTIQ	360
DB	296	VNMASMNWYNNAPPSAIEAAAIRSPHLLDFLEQLTIFSASSRWSNTRMTYWRGHTIQ	355
QY	361	SRPTGGGLNTSTHGSTNTSINPVRLSPFSDRVYVWTSYAGVLLWGIYLEPHGVPTVREN	420
DB	356	SRPTGGGLNTSTHGATNTSINPVTLPFASRDVYVWTSYAGVLLWGIYLEPHGVPTVRFN	415

QY 421 FRNPONTFERGTANYSPYSPGLQKDSLETLPETTERPNVESYSHRLSHIGLISQSR 480
 DB 416 FTNPONISDRGTANYSPYSPGLQKDSLETLPETTERPNVESYSHRLSHIGLISQSR 475
 QY 481 VHPVYVSWTHRSADRTNTISDSITQPLVKSFNLNSGTSVSGPGTGGDIIRTNVNGS 540
 DB 476 VNPVYVSWTHRSADRTNTIGPNRTQIPMKVASELPQGTIVRPGTGGDILLRNTTGG 535
 QY 541 VLSGLNPNNTSLQRYRVRYAASQTMVLRLVTVGSTTFDQGPPTMSANESLTSOSPR 600
 DB 536 FGIIRVTVNGPLQRYRIGRYASTVDFDFVSRGGTIVNFRFLRTMWSGDELKYGNFV 595
 QY 601 FAREPVGISAGSGQ-TAGISISNAGQTFHPKIEPIPTATFEAYDYDLERAQEAVAL 659
 DB 596 RRAFTPTFTIQDIIRTSIQGLSGNGEYTIKIEIIPVATFEAYDYDLERAQEAVAL 655
 QY 660 FTNINPRLKTDVTDYHIDQVSNLVACLSDEFCLDEKRELEKVKYAKRLSDERNLLQDP 719
 DB 656 FTNINPRLKTDVTDYHIDQVSNLVACLSDEFCLDEKRELEKVKYAKRLSDERNLLQDP 715
 QY 720 NFTSINKQDPFISTNEOSNFTSIHQSEHGWSGSENIITIOGNDVFKENYVTLPGTFNEC 779
 DB 716 NFTSINKQDPFISTNEOSNFTSIHQSEHGWSGSENIITIOGNDVFKENYVTLPGTFNEC 775
 QY 780 YPTLYOKIGESL-KAVTRYOLRGYIEDSDOLLEIYLRYNKAKHETLDVPGTESVWPLSVE 839
 DB 776 YPTLYOKIGESL-KAVTRYOLRGYIEDSDOLLEIYLRYNKAKHETLDVPGTESVWPLSVE 835
 QY 840 SPIGRGCEPNRCAPHFENWPDLCSDRGKCAHSHHFSLDIDIGCTDLHENLGVWVVF 899
 DB 836 SPIGRGCEPNRCAPHFENWPDLCSDRGKCAHSHHFSLDIDIGCTDLHENLGVWVVF 895
 QY 900 KIKTOEGHARLGNLEFIEEPLGELSRVKRAEKKWRDKREKLOLETKRYVYTAKEAVD 959
 DB 896 KIKTOEGHARLGNLEFIEEPLGELSRVKRAEKKWRDKREKLOLETKRYVYTAKEAVD 955
 QY 960 ALFVDSQYDLQADTNGIMHAADKLVHRIEAYLSLSVPGVNAEIPFELSGRIITAI 1019
 DB 956 ALFVDSQYDLQADTNGIMHAADKLVHRIEAYLSLSVPGVNAEIPFELSGRIITAI 1015
 QY 1020 SLYDARNVVKNGDFNGLACVNVKGVHDVQOSSHRSVLVPIPEWAEVQAVRVCPCGYI 1079
 DB 1016 SLYDARNVVKNGDFNGLTCWNVKGHDVQOSSHRSVLVPIPEWAEVQAVRVCPCGYI 1075
 QY 1080 LRVYAYKEGEGCVTHIENNTDELKFKNCEEEVYPTDGTCTNDYTAHQGTA---V 1135
 DB 1076 LRVYAYKEGEGCVTHIENNTDELKFKNCEEEVYPTDGTCTNDYTAHQGTA---V 1135
 QY 1136 CNSRNAGYEDAYEVDTTASVNYKPTVEEYTYDVRDNDHCEYDRGVYVNPPLPAGYMTKE 1195
 DB 1136 CNSRNAGYEDAYEVDTTASVNYKPTVEEYTYDVRDNDHCEYDRGVYVNPPLPAGYMTKE 1195
 QY 1196 LEYFPETDKVIBIGETEGKFIIVDSVELLMEE 1228
 DB 1196 LEYFPETDKVIBIGETEGKFIIVDSVELLMEE 1228

RESULT 2
 A48970
 parasporal crystal protein cryIcB - Bacillus thuringiensis
 C;Species: Bacillus thuringiensis
 C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C;Accession: A48970
 R;Kaiman, S.; Kiehne, K.L.; Libs, J.L.; Yamamoto, T.
 Appl. Environ. Microbiol. 59, 1131-1137, 1993
 A;Title: Cloning of a novel cryIC-type gene from a strain of Bacillus thuringiensis subs
 A;Reference number: A48970; MUID:93236401; PMID:8476286
 A;Accession: A48970
 A;Status: preliminary
 A;Molecule type: nucleic acid
 A;Residues: 1-1176 <KAL>
 A;Cross-references: UNIPROT:P56953; GB:M97880; NID:g2899263

A;Experimental source: subsp. galleriae HD29
 A;Note: sequence extracted from NCBI backbone (NCBIN:129672, NCBIIP:129675)
 C;Superfamily: parasporal crystal protein
 C;Keywords: delta-endotoxin
 Query Match 58.5%; Score 3791; DB 2; Length 1176;
 Best Local Similarity 62.3%; Pred. No. 3.9e-233;
 Matches 762; Conservative 125; Mismatches 241; Indels 96; Gaps 16;
 QY 40 CIAAGNNINPLVASVTVQG---INTAGRLGLV---GVPPAGQIASFVSFLVGLGLWPRG 93
 DB 14 CLSNPEI--LLDGERISTGSSSIDSLVQLLVGNFVPGGGLVGLLDFVWGVIGVP-- 69
 QY 94 RDOWEFLHEVEQLINQITENARNTALARLOGLGDSFRAYQOSLSDLENRDRDARTSV 153
 DB 70 -SPWDAFLVQIEILINERIAAARSAAISNLEGLGNFNFIYVEAFKEWADPNPVTIR 128
 QY 154 LYQYTALELDLFNANPLFAIRNQEVPLLMVYAAANLHLLLRDASLFGSEGLTSQEI 213
 DB 129 VVDRFRILQGLLERDIPSPRIAGFEVPLLSVYAAANLHLLLRDSSIFGARWGLTINV 188
 QY 214 QRYERQVQTRDYDYCVWEYNTGLNSLRGTNAASWRYNQPRDLTLGLVLDLVALFPS 273
 DB 189 NENYNELIRHIDEYANHCADTYNRGLNLPKSTYQDWITYNRLRLDLTLVLDIAAFAFPS 248
 QY 274 YDRTYPIINTSAQLTRVYTDALGATGVMAWNNWNNAPSFSALETAVIRSPHLLDPL 333
 DB 249 YDNRPIQSVQGLTREIYTDPLITPNQLOSV---AQLPTFNWESNAIRTPHLDVL 304
 QY 334 EQTIESTSRMSATRHMTYWRGHTIQSRPIGGLNTSTHGSTNTSINPVLRSFFSRDVI 393
 DB 305 NNLTITFD---WFSGRNFVWGHRVYSNRIGGNTSPYIGREANQEPSPRFTNGPVF 361
 QY 394 WTESYAGVLLWGLYLPPIHGVPTVR-----FN-----FRNPTQTE---RGT 432
 DB 362 RTLSN-----PTFRLQPPWAPPENLRGVEGVEFSTPLNSFTYRGRGT 405
 QY 433 ANYSQPVESPGQLKDSLETLPETTERPNVESYSHRLSHIGLISQSRVHV---PVYS 488
 DB 406 V-----DSLTELPPEDNSVPREYSHRLCHATFVQSRGTFFLTGPFVFS 451
 QY 489 THRSADRTNTISDSITQPLVKSFNLNSGTSVSGPGTGGDIIRTNVNGSVLSMGLNF 548
 DB 452 THRSADRTNIIIPDVINQIPLVKAPNLTSQTSVVRGPGTGGDIIRTNVNGSVLSMGLNF 511
 QY 549 NNTSLQRYRVRYAASQTMVLRLVTVGSTTFDQGPPTMSANESLTSOSFRAEPVGI 608
 DB 512 SNTTLQRYRVRYAASQTMVMSVTVGGSTTGNQGPPTMSANGALTSOSFRAEPVGI 571
 QY 609 SASGQTAGISISNAGROTFFDKLEFPIATPEAYDLERAQEAVALFTNTNPRRL 668
 DB 572 SASGQCASISISNNVGRQWFLDRIEFLPVSTFEEYDLERAQEAVALFTNTNQLGL 631
 QY 669 KTDVTDYHIDQVSNLVACLSDEFCLDEKRELEKVKYAKRLSDERNLLQDPNFTSINKQ 728
 DB 632 KTDVTDYHIDQVSNLVACLSDEFCLDEKRELEKVKYAKRLSDERNLLQDPNFTSINKQ 691
 QY 729 DFISTNEQSNFTSIHQSEHGWSGSENIITIOGNDVFKENYVTLPGTFNECYTYLYOKI 788
 DB 692 D-----RGWRGSDTITIQGDDVFKENYVTLPGTFNECYTYLYOKI 733
 QY 789 GSELKAYTRYOLRGYIEDSDOLLEIYLRYNKAKHETLDVPGTESVWPLSVPESPIGRGCEP 848
 DB 734 DESKLKSYTRYELRGYIEDSDOLLEIYLRYNKAKHETLDVPGTESVWPLSVPESPIGRGCEP 793
 QY 849 NRCAPHEWNPDLDCSDRGKCAHSHHFSLDIDIGCTDLHENLGVWVVFVFKITQEGHA 908
 DB 794 NRCAPHEWNPDLDCSDRGKCAHSHHFSLDIDIGCTDLHENLGVWVVFVFKITQEGHA 853
 QY 909 RIGNLEFIEEKLIGELSRVKRAEKKWRDKREKLOLETKRYVYTAKEAVDALFVDSOYN 968
 DB 854 RIGNLEFIEEKLIGELSRVKRAEKKWRDKREKLOLETKRYVYTAKEAVDALFVDSOYN 913

Qy 969 RLQADTNIGMHAADKLVRIRIAYLSLSVLPQVNAEIEELEGRIITAIISLYDARNVV 1028
 Db 914 RLQADTNIAIHAADKRVHRIAYLPELSVLPQVNAEIEELEGRIITAIISLYDARNVI 973
 Qy 1029 KNGDFNGLACNNVKGHDVY-QQSHHRSVLVPEWEAEVQAVRVCPCRGYILRVATYKE 1087
 Db 974 KNGDFNGLACNNVKGHDVY-QQSHHRSVLVPEWEAEVQAVRVCPCRGYILRVATYKE 1033
 Qy 1088 GYEGCVTIHEIENNTDELKPKCEBEVYPTDTGTCTNDVTAHQ---GTAVCNMRNAGYE 1144
 Db 1034 GYEGCVTIHEIENNTDELKPKCEBEVYPTDTGTCTNDVTAHQ---GTAVCNMRNAGYE 1092
 Qy 1145 DAYEVDITASVNYKPTVEETTYDVRDNHCEYDGRGVNYPPLPAGVMTKELEYFPETDK 1204
 Db 1093 KSYESNSQADVASVEYKADYDGRDNHCEYDGRGVNYPPLPAGVMTKELEYFPETDK 1152
 Qy 1205 VWIEIGETEGKFIIVDSVELLMEE 1228
 Db 1153 VWIEIGETEGKFIIVDSVELLMEE 1176

RESULT 3
 S00944
 Parasporal crystal protein cry1Cal - Bacillus thuringiensis (strain entomocidus 60.5)
 C;Species: Bacillus thuringiensis
 C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
 C;Accession: S00944
 R;Honee, G.; van der Salm, T.; Visser, B.
 Nucleic Acids Res. 16, 6240, 1988
 A;Title: Nucleotide sequence of crystal protein gene isolated from B. thuringiensis subsp. entomocidus
 A;Reference number: S00944; MUID:88289380; PMID:3399402
 A;Accession: S00944
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-1169 <HON>
 A;Cross-references: UNIPROT:P05518; EMBL:X07518; NID:g40293; PIDN:CAA30396.1; PID:g40294
 C;Superfamily: parasporal crystal protein
 C;Keywords: delta-endotoxin

Query Match 53.8%; Score 3484.5; DB 2; Length 1189;
 Best Local Similarity 56.8%; Pred. No. 1.4e-213;
 Matches 715; Conservative 143; Mismatches 291; Indels 109; Gaps 20;

Qy 7 NENEII--NALSTPAVNSHSTQMDLSPARIESLCIAEGNNINPLVASSTVQTGINIAG 64
 Db 5 NQNCIPYNCLS-----NPEVLLDGERISTGN-----SSIDISLSLVQ 43
 Qy 65 RIIGVLGVPAGQIASFYSFLVGLWPRGRDOWEIFLEHVEQLINQOITENARTALRL 124
 Db 44 FLVSNF-VPGGFLVGLIDFVWGVGP---SQMDAFLVQIEQLINERIAEFARNAAIL 99
 Qy 125 QGLGDSFRAYQOGLDLENRDDARTSVLYTVQIALELOFLNAMPULFARNQOEVLPMV 184
 Db 100 EGLONNFNIYFAKWEEDPNPETHRVIDRILDLGLLDRIDPSFRISGEVPLLSV 159
 Qy 185 YAAQANLHLLLRDASLFGSEFGTSGEIQRYXERQVEQTRDYSQVCEWYNTGLNSLRG 244
 Db 160 YAAQANLHLLLRDASLFGSEFGTSGEIQRYXERQVEQTRDYSQVCEWYNTGLNSLRG 219
 Qy 245 TNAASVRYNQFRDLTLGLVDLVALPSPYDTRTYPTINTSAQLTREYVTDAGATGVNMA 304
 Db 220 STYQDMWYNYLRDLTLGLVDLVALPSPYDTRTYPTINTSAQLTREYVTDAGATGVNMA 275
 Qy 305 SMWYNNNAPSFAIETAVIRSHLLDFLEQLITFSTSRWSATRHMTYWRGTIOGRPI 364
 Db 276 PQIQSAQLPFTFNWESSRIRNPHLFDILNLTFTD---WFSVGRNRYWGGHRISSLI 332
 Qy 365 GGGINTSTHGSNTSINPVLSPFSRDVYWTESYAGVLL---WGIYLEPIHGVTYFRN 420
 Db 333 GGGNITSPIYGRANQEPSPFSFNGVFTLNSPTLRLQWPAPPPNLRGVEGVEFS 392
 Qy 421 FRNPQNTFE---RGTANYSQPYSPGLQLKDSLETLPETTERPNYESYSHRLSHIGLS 477

Db 393 --TPTNSFTYRGRGTV-----DSLTELPPEDNSVPPREGYSHRLCHATFVQ 436
 Qy 478 QSRVHV-----PVYSWTHRSADRTNTISSDITQIPLVKSFNLSNGTSSVSGPGTGGDII 533
 Db 437 RSGTPFLTITGWFSWTDRTSATLNTIDPERINQIPLVKGFRVMGCTSVITGPGTGGDIL 496
 Qy 534 RTNVGSLVSMGLNFNNTSLQRYRVRVRYAASQ-----TMVLRTVTVGGSTTFDQGFPS 586
 Db 497 RRNTFGDFVSLQVNIINSPIQRYRLFRYASSRDARVILVTGAASTGVGGQVSVNMPLOK 556
 Qy 587 TMSANESLTSQSFRFAEFP-----VGISASGSQTAGISISNNAAGRTQTFPHFKIE 635
 Db 557 TMEIGENLTSRTYTDSPNPFPRANDIIGISEQLFCAG-SISGS-----ELYIDKIE 611
 Qy 636 FIPITATFEAYDLERAQEAVALFTNPNRRLKTDVTDYHIDQVSNLVACLSDSEFCLE 695
 Db 612 IILADATFEAESDLERAQKAVNALFTSSNQIGLKTVDVTDYHIDQVSNLVACLSDSEFCLE 671
 Qy 696 KRELLEKVKAKELSDERNLLQDPNFTSINKQPDFISTNEQSNTFSIHEQSEHGHWGSEN 755
 Db 672 KRELSEKVKHAKELSDERNLLQDPNFRGINRQPD-----RGWRGSTD 713
 Qy 756 ITIQEGNDVPKENYVTLPGTFNECYPTLYQKIGESLKAATRYQYLRGYIEDSODLIYL 815
 Db 714 ITIQGGDDVPKENYVTLPGTVDECYPTLYQKIDESKLKATRYELRGYIEDSODLIYL 773
 Qy 816 IRYNAKHETLDVFGTSVMPLSVESPIGRCEGPNRCAPHFEWNPDLDCSCRDGKCAHHS 875
 Db 774 IRYNAKHETLDVFGTSVMPLSVESPIGRCEGPNRCAPHFEWNPDLDCSCRDGKCAHHS 833
 Qy 876 HHSFLDIDGCTDLHNLGLVWVVFVKIQDGHARLGNLEFIEKPLLGELALSVRKAEKK 935
 Db 834 HFTDLDVGTDLNEDLGVVVFVKIQDGHARLGNLEFIEKPLLGELALSVRKAEKK 893
 Qy 936 WRDKREKLETKRVYVTEAKEDVADLVDSQYNRLOADTNIGMHAADKLVRIRIAYLS 995
 Db 894 WRDKREKLETKRVYVTEAKEDVADLVDSQYNRLOADTNIGMHAADKLVRIRIAYLS 953
 Qy 996 ELSVIFGVNAEIEELEGRIITAIISLYDARNVKNKGFNNGLACNNVKGHDVY-QQSHHR 1054
 Db 954 ELSVIFGVNAEIEELEGRIITAIISLYDARNVKNKGFNNGLACNNVKGHDVY-QQSHHR 1013
 Qy 1055 SVLVPEWEAEVQAVRVCPCRGYILRVATYKEGEGCVTIHEIENNTDELKPKCEEE 1114
 Db 1014 SVLVPEWEAEVQAVRVCPCRGYILRVATYKEGEGCVTIHEIENNTDELKPKCEEE 1073
 Qy 1115 EYVPTDTGTCTNDVTA---HOGTAVCNMRNAGYEDAVEDVTASVNYKPTVEETTYDVR 1170
 Db 1074 EYVPTDTGTCTNDVTA---HOGTAVCNMRNAGYEDAVEDVTASVNYKPTVEETTYDVR 1131
 Qy 1171 RDNHCEYDGRGVNYPPLPAGVMTKELEYFPETDKVMIIEGTEGKFIIVDSVELLMEE 1228
 Db 1132 RENPCSNRGYDGYTLPAGVMTKELEYFPETDKVMIIEGTEGKFIIVDSVELLMEE 1189

RESULT 4
 S11446
 parasporal crystal protein cry1Dai - Bacillus thuringiensis
 N;Alternate names: parasporal crystal protein cry1D
 C;Species: Bacillus thuringiensis
 C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 C;Accession: S11446
 R;Hoefte, H.; Soetaert, P.; Janssens, S.; Peferoen, M.
 Nucleic Acids Res. 18, 5545, 1990
 A;Title: Nucleotide sequence and deduced amino acid sequence of a new Lepidoptera-specific
 A;Reference number: S11446; MUID:91016842; PMID:2216728
 A;Accession: S11446
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1165 <HOE>
 A;Cross-references: UNIPROT:P19415; EMBL:X54160; NID:g40279; PIDN:CAA38099.1; PID:g40280
 C;Superfamily: parasporal crystal protein
 C;Keywords: delta-endotoxin

QY 581 DQGFPPSTMSANESLTSOSFFAFPPVGVISASGSQTA-----GISISNNAGROTTFHFDKI 634
 DB 533 IGNFSSMTSGDDLQYGRFRFVAGFTTPTFSDANSTFTTGAFFSPNNEV-----YIDRI 587
 QY 635 EFTPTATFAEYDLERAQAVNALFTNTNPRKLTDTVDYHIDOVSNLVACLSDFECLD 694
 DB 588 EFPFAEYTFPAEYDLERAQAVNALFTSSNQIGLTKTDTVDYHIDKVSNNLVECLSDSFCLD 647
 QY 695 EKRELEKVKYAKRLSDERNLLQDPNFTSINKOPDFISTNEQSNFTSIHQSEHGHWGSE 754
 DB 648 EKRELEKVKYAKRLSDERNLLQDPNFRGNRPD-----RWRGST 699
 QY 755 NITIQEONDVFNKENVYTLPTFTNECYPTIYKIGSESELKAYTRYQLRGYIEDSQDLEIY 814
 DB 690 DITIQGDDVDVFNKENVYTLPTFTGCGYPTIYKIDESKLVYTRYQLRGYIEDSQDLEIY 749
 QY 815 LIRYNKHEITLDVPGTESVMPLESPIGRCGPNRCAPHFENPDLDCCSCRGEKCAHH 874
 DB 750 LIRYNKHEITVNVPGTSGSLMPLSAQSPIGKCGPNRCAPHLEWNPDLDCSCRGEKCAHH 809
 QY 875 SHFSLDIDIGCTDLHENLGVVWVFKIKTOEGHARLGNLFEBEKLPLGALGRVKAER 934
 DB 810 SHFSLDIDIGCTDLNEDLGWVFKIKTOEGHARLGNLFEBEKLPLGALGRVKAER 869
 QY 935 KWRDKREKLETKRVVYTRAKEAVDALFVDSQYRNLQADTNIGMHAADKLVHRIRAYL 994
 DB 870 KWRDKREKLETNIVYKEAKESVDALFVNSQYDQLQADTNIAMHAADKRVHSIRAYL 929
 QY 995 SELSVIRPGVNAEYFEELEGRITAIISLYDARNVKNCGDNGLACNVKGVHDV-QQSHH 1053
 DB 930 PELSVIRPGVNAEYFEELEGRIFTAFSLYDARNVKNCGDNGLACNVKGVHDVVEQNNH 989
 QY 1054 RSVLVPWEAEYSQAVRVCGRGYILRVYAYKEGEGCVTHIEINNTDELKFKNCEE 1113
 DB 990 RSVLVPWEAEYSQAVRVCGRGYILRVYAYKEGEGCVTHIEINNTDELKFKNCVE 1049
 QY 1114 EYVYPTDTGTCNDYTAHQG--TAVCSNRNAGYEDAYEVDVTASVNYKPYTEETTYDVR 1171
 DB 1050 EYVYPTVTCNDYTANQBEYKAYTSHNRGYDEAYGNPNPVPADYVYEEKAYTDGR 1109
 QY 1172 DNHECEVDRGVNYPPLPAGYMTKELEYFPETDKWIEIGTEGKFIVDSVELLIMEE 1228
 DB 1110 ENPCESNRGYDTPLPAGYMTKELEYFPETDKWIEIGTEGKFIVDSVELLIMEE 1166

RESULT 6
 A41052
 parasporeal crystal protein cryAel - Bacillus thuringiensis (strain alesti)
 C;Species: Bacillus thuringiensis
 C;Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 09-Jul-2004
 C;Accession: A41052
 R;Lee, C.S.; Aronson, A.I.
 J. Bacteriol. 173, 6635-6638, 1991
 A;Title: Cloning and analysis of delta-endotoxin genes from Bacillus thuringiensis subsp
 A;Reference number: A41052; MUID:92011442; PMID:1655719
 A;Accession: A41052
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1181 <LEE>
 A;Cross-references: UNIPROT:Q03748; GB:M65252; NID:g142874; PIDN:AAA22410.1; PID:g142875
 C;Superfamily: parasporeal crystal protein
 C;Keywords: delta-endotoxin

Query Match 52.98; Score 3426.5; DB 2; Length 1181;
 Best Local Similarity 56.94; Pred. No. 6.7e-210;
 Matches 700; Conservative 148; Mismatches 304; Indels 79; Gaps 22;

QY 27 MDLSPDARDISLCAEGNNINP---LVASASTVQTG---INIAGRILGLV---GVPPAGQ 77
 DB 1 MDNNP--KINE--CIPYCNLSNPEVEVEGERIETGTPIDISLSLTQFLLSEFVPGAGF 56
 QY 78 IASFYSLVGLMWRGRDOWEIFLEHVEQLINQIQITENARNTALRLOGLGDSFRAYQOS 137

DB 57 VLGLIDLWGFVGP---SQMDAFLVQLEQLISQRIEIEFARNQAIISLEGUSNLQIYAE 113
 QY 138 LEDWLENRDDARTRSVLYTYQYIALELDFNAMPLFAIRNOEVEPLMYAAQANLHLLLR 197
 DB 114 FREADPTNPALREEMRIQFNDMSALTTAIPLFTVQNVQVPLLSYVQAVNLHLSVL 173
 QY 198 DASLFGSEFGLTSQETQRYRVEROQTRDYSDYCVIEWYNTGLNSLRGTNAASVRYNQFR 257
 DB 174 DVSVFQRMGLDVATINSRYNDLTRIGTYTDAVRMYNTGLERVGMGPDSDRDVRYNQFR 233
 QY 258 RDLTLGLVDLVALFSPDYTRTYPINTSAQLTRVYTDATGATGVNMAVMWYNNAPSF 317
 DB 234 REUTLVLDIVSLFPNYSRTYPIRTVSQUTREIYT-----NPVLENFDSFRGSAQ 285
 QY 318 AIEATVIRSHPLDLFLEQLTIFSTSRWSATRMITYWRGHTIOSRPITGGG---LNTSTHG 374
 DB 286 RIFQS--IRSPHLDILNLSITYT---DAHGGVYVWSGHQIMASPVGSGPEPTPLYG 339
 QY 375 STWTSINPVL--SFFRDRDVWTES---YAGVLWGIYLEPIHGVTVRFNRPQNTFER 430
 DB 340 TMGNAAPQQRIVAQLGQGVYRTLSSTFYRNPFTIIGINNORLSVLDGTFEYAGSSNL--- 396
 QY 431 GTANYSOPYESPGLQKDSLETLPETTERPNYESYSHRLSHGLI-----SQSRVHV 483
 DB 397 ----PSAVRKSG--TVDSLDEIPQDNNVPPQGFGRSLRSHVSMFRSGFSNVSIVIRA 450
 QY 484 PVYSWTHRSADRNTWTISSDSITQIPLVKSFNLSGTSVSVSGPGFTGGDIIRTNVNGSVLS 543
 DB 451 PMFSWIRHSABFNIIIPSSQITQIPLTKSTNLGSGTSVVKSGPGFTGGDILRRTPSGQIST 510
 QY 544 MGLNFNTLSQRYRVRYVAASQTMVLRVTVGSTTFDQGFPPSTMSANESLTSQSFRFAE 603
 DB 511 LRVNITAPLSQRYRVRYVAASQTMVLRVTVGSTTFDQGFPPSTMSANESLTSQSFRFAE 570
 QY 604 FPGVIS-ASQSQTAGTISIS--NNAGRTFHPDKTEFPTATFAEYDLERAQAVNALFT 661
 DB 571 FTTPFPFSGSSVFTLSAHVFNSENVY--IDREFVPAEVTFAEYDLERAQAVNALFT 629
 QY 662 NTPRELKTDVTDYHIDOVSNLVACLSDFECLDEKRELEKVKYAKRLSDERNLLQDPNF 721
 DB 630 SPNQIGLTKTDTVDYHIDOVSNLVACLSDFECLDEKRELEKVKYAKRLSDERNLLQDPNF 689
 QY 722 TSINKOPDFISTNEQSNFTSIHQSEHGHWGSENIITQEGNDVFNKENVYTLPTGTFNFCYP 781
 DB 690 RGINRQPD-----RWRGSTDITIQGDDVDVFNKENVYTLPTGTFNFCYP 731
 QY 782 TYLYQKIGSESELKAYTRYQLRGYIEDSQDLEIYLRYNKHEITLDVPGTESVMPLESV 841
 DB 732 TYLYQKIGSESELKAYTRYQLRGYIEDSQDLEIYLRYNKHEITVNVPGTSGLMPLSFESS 791
 QY 842 IGRGSPNRCAPHFENWNPDLDCSCRGEKCAHSHHPSLDDIGCTDLHENLGVVWVFKI 901
 DB 792 IGRGSPNRCAPHFENWNPDLDCSCRGEKCAHSHHPSLDDIGCTDLHENLGVVWVFKI 851
 QY 902 KTOEGHARLGNLFEBEKLPLGALGRVKAERKWRDKREKLETKRVYTRAKEAVDAL 961
 DB 852 KTOEGHARLGNLFEBEKLPLGALGRVKAERKWRDKREKLETKRVYTRAKEAVDAL 911
 QY 962 FVDSQYRNLQADTNIGMHAADKLVHRIRAYLSLSVIFGVNAEYFEELEGRITAIISL 1021
 DB 912 FVNSQYDQLQADTNIAMITADKRVHRIQYAYLPELSVIFGVNAGIFEELEGRIFTAYSL 971
 QY 1022 YDARNVKNCGDNGLACNVKGVHDV-QQSHRSVLVPIPEWAEYSQAVRVCGRGYIL 1080
 DB 972 YDARNVKNCGDNGLACNVKGVHDVVEEQNNHRSVLVPEWAEYSQAVRVCGRGYIL 1031
 QY 1081 RVTAYKEGEGCVTHIEINNTDELKFKNCEEVYPTDTGTCNDYTAHQG---GTAVCN 1137
 DB 1032 RVTAYKEGEGCVTHIEINNTDELKFKNCEEVYPTDTGTCNDYTAHQG---GTAVCN 1090
 QY 1138 SRNAGYEDAYEVDVTASVNYKPYTEETTYDVRDNHCEYDRGVNYPPLPAGYMTKELE 1197

578 RIAEELPIRGELYIDKIELILADATFEESYDLERAQAVNALFTSTNQGLKTDVTDYH 633
 677 IDQVSNLVACLSDFCFLDEKRELLEKVKYAKRLSDERNILQDPNFTSINKQDPDISTNEQ 736
 638 IDQVSNLVECLSDFCFLDEKRELSEKVKHAKRLSDERNILQDPNFRGINQDP----- 690
 737 SNFTSIHSESHGHWGSENIITIQEGNDVFKENYVTLPGTTFNECYPYLYQKIGSESLKAY 796
 691 -----RCWGRGSTDITIQQGDVDFKENYVTLPGTTFNECYPYLYQKIDSKLKAY 739
 797 TRYQLRGYIEDSDIATYILIRYNNAKETHLDVPGTESWPLSVSPICRCEPNRCAPHFE 856
 740 TRYELRGYIEDSDIATYILIRYNNAKETHVNPVGTSLWPLSAOSPIGKCGEPNRCAPHLE 799
 857 WNPDLDCSCRDGKCAHSHHFSLDIDIGCTDLHENLGVWVFKIKTOEGHARLGNLEFI 916
 800 WNPDLDCSCRDGKCAHSHHFSLDIDVGCTDLNEDLGWVWIFKIKTQDGVARLGNLEFL 859
 917 EEKPLLGEALSRVKRAEKKWRDKREKLQLETKRVTYEAKEAVDALFVDSQYNRLQADTNI 978
 860 EENPLLGEALSRVKRAEKKWRDKCELEWETNIVYKEAKESVDALFVNSQYDRLQADTNI 919
 977 GMIHAADKLVRITREAYLSLSVIGPVNABIFPELEGRIITALISLYDARNVVKNGDFNNG 1036
 920 AMTHAADKRVHSITREAYLPESLVIPIGVNAAIFPELEGRIPTAFSLYDARNVVKNGDFNNG 979
 1037 LACWNVKGHDV-QQSHHRSVLVIPSEWAEVSQAVRCPGRGYILIRVTAKEGYGECGVT 1095
 980 LSCWNVKGHDVVEEQNNHRSVLVPPSEWAEVSQEVRCVPCRGYILIRVTAKEGYGECGVT 1039
 1096 IHEIENNTDELFPKNCBEEVYPTDTGTGNDYTA----HOGTAVCNRSNAGYEDAYEVD 1151
 1040 IHEIENNTDELFPKNCBEEVYPTDTGNTYTAQEEHGT--YTSRNGYDEAYESNS 1097
 1152 TASVNTKPYEETTYDVRDNHCEYDVRGVNVPPLPAGVMTKELEYFPPTDKVWIBIGE 1211
 1098 SV---HASVYEEKSYDTRRENPCESNRGVDYTPLPAGVYTKLELEYFPPTDKVWIBIGE 1154
 1212 TGKFTIVDSVELLMEE 1228
 1155 TGTFTIVDSVELLMEE 1171

 RESULT 8
 S32649
 parasporal crystal protein cryIFa3 - *Bacillus thuringiensis*
 C:Species: *Bacillus thuringiensis*
 C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
 C:Accession: S32649
 R:Lambert, B.
 submitted to the EMBL Data Library, April 1993
 A:Reference number: S32645
 A:Accession: S32649
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1174 <LAW>
 A:Cross-references: UNIPROT:Q45749; EMBL:Z22512; NID:g295865; PIDN:CAA80235.1;
 C:Superfamily: parasporal crystal protein
 C:Keywords: delta-endotoxin

 Query Match 52.7%; Score 3413.5; DB 2; Length 1174;
 Best Local Similarity 57.3%; Pred. No. 4.5e-209;
 Matches 685; Conservative 139; Mismatches 264; Indels 107; Gaps 19;

 72 VPFAGQIASFYSLVGLMWRGRDQWEIFLEHVQQLINQIITENANTALARIQGLGDSF 131
 49 VPGVGVAFLDPLIWGITFP---SEWSLFLQTEQLIEQRIETLERNRAITTLRGLADSY 105
 132 RAYQOSLEDWLENRDARFERSVLYTQYIALEDLFLNAMPLEFAIRNOEVLPLVYQAANL 191
 106 EYVLEALREWEENPNNAQRLDVRIRFANTD DALITAINNFTLSITSEIPLLSVYVQAANL 165
 192 HILLRLDASLFGSEFGLTSQETQRYVYERQVEOTRDSYCVIEWYNTGLNSLRGTNAASWV 251

Db 681 NLLQDPNFRGINRQLD-----RGMWGSDTITIQGDDVFKENYVTL 722

Qy 774 GTFNECYPTLYOKIGESKAYTRYQYLRGYIEDSDLEIYLIRYNAKHETLDVPGTESV 833

Db 723 GTFDECTPTLYOKIDESKAYTRYQYLRGYIEDSDLEIYLIRYNAKHETVNVPGTGS 782

Qy 834 WPLSVESPIGRGPNRCAPHEFNWPDLDSCRDGKCAHSHHFSLDIDIGCTDLHENL 893

Db 783 WPLSAQSPIGKGPENRCAPHEFNWPDLDSCRDGKCAHSHHFSLDIDVGCCTDLNEDL 842

Qy 894 GVVVFKIKTQEGHARLGNLFIEBKPLGLBALSRVRAEKAKKWDKREKLOLETKRYVTE 953

Db 843 GVVVFKIKTQEGHARLGNLFIEBKPLGLBALSRVRAEKAKKWDKREKLEWETNIVKE 902

Qy 954 AKEAVDALFVDSQYRLQADTNIGMHAADKLVRHREAYLSLSLPGVNAIFEELEG 1013

Db 903 AKESVDALFVNSQYDQLQADTNIGMHAADKRVHSIREAYLPELSVIPGVNAALFEELEG 962

Qy 1014 RIITAFSLYDARNVYKRGDFNNGLACVNVKGVHDV-QQSHRSVLVPIPEWAEVSAQVRV 1072

Db 963 RIITAFSLYDARNVYKRGDFNNGLACVNVKGVHDVQSHRSVLVPIPEWAEVSAQVRV 1022

Qy 1073 CPGRGYILRVYAYKEGEGCVTHIEINNTDELKFKNCBEEVYPTDGTCDNDYTAHQ- 1131

Db 1023 CPGRGYILRVYAYKEGEGCVTHIEINNTDELKFKNCBEEVYPTDGTCDNDYTAHQ- 1082

Qy 1132 --GTAVCNRSNAGYEDAYEDVTTASVNVKPYEETVTDVRRDNHCEYDRGVNYPPLPA 1189

Db 1083 EYGGAY-TSRNGVNEAPSV----PADYASVYEKSYTDORRENPCBFNGRYDTPPLV 1137

Qy 1190 GYMTKELEYFPETDKVWIEIGETGKFIIVDSVELLMEE 1228

Db 1138 GYMTKELEYFPETDKVWIEIGETGKFIIVDSVELLMEE 1176

RESULT 10

A42459

Parasporal crystal protein cryIa1 - Bacillus thuringiensis (strain aizawai)

N;Alternate names: parasporal crystal protein cryIa

C;Species: Bacillus thuringiensis

C;Date: 10-Jul-1992 #sequence_revision 10-Jul-1992 #text_change 09-Jul-2004

C;Accession: A42459

R;Chambers, J.A.; Jelen, A.; Gilbert, M.P.; Jany, C.S.; Johnson, T.B.; Gawron-Burke, C.

J. Bacteriol. 173, 3966-3976, 1991

A;Title: Isolation and characterization of a novel insecticidal crystal protein gene fcd

A;Reference number: A42459; MUID:91286178; PMID:2061280

A;Accession: A42459

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1174 <CHA>

A;Cross-references: UNIPROT:Q03746; GB:M63897; NID:g142757; PIDN:AAA22348.1; PID:g142758

C;Superfamily: parasporal crystal protein

C;Keywords: delta-endotoxin

Query Match 52.7%; Score 3412.5; DB 2; Length 1174;

Best Local Similarity 58.4%; Pred. No. 5.2e-209;

Matches 693; Conservative 132; Mismatches 271; Indels 91; Gaps 21;

Qy 72 VPFAGQIASFYSFLVGLWPRGRDWEIIEHVEQLINQOITENARTALRLOGLDGSF 131

Db 49 VPGVGVAFLDILWGFIIP--SDSLFLLQIEQLIEQRIETLERNRATITLUGLADSY 105

Qy 132 RAYQQSLDLENRDARTSRVLYTQYIALEDFLNAPLFAIRNOEVPMLLVYQAANL 191

Db 106 EIVIEALREWEANPNNAQLREDVIRPANTDDALITAINNFTLTSFEIPLLSVYQAANL 165

Qy 192 HLLLRDASLFGSEFGLTSQEIQRYERQVEQTRDYSYDCEWYNTGLNSLRGTNAASV 251

Db 166 HLLSLRDVAFSGQGLDIAVNVHNRNLNLRHYTKHCLDTYNQGLNLRGTNRQWA 225

Qy 252 RYNOQRDLTLGLVLDLVALPPSDTRYTPINTSAQLTREYVYDIAIGATGVNMAWMWYNN 311

Db 226 RYNOQRDLTLGLVLDLVALPNDVRYTPYQTSQLTREYVYDIAIGATGVNMAWMWYNN 278

Qy 312 NAPS-PSAIAETAVIRSPHLLDFLEQLTIFSTSRWSATRHMTYWRGHTIOSRPIGGG-LN 369

Db 279 NIPNGFNRAEFGV-RPHLMDFMNSL--FVTA---ETVRSQTWGGHLSVSRRTAGNRIN 332

Qy 370 TSHGSDTNTINPVRSLFFS----RDVWYTESYAGVLLMGVILEPHTIGVPTVAFNFRNPQ 425

Db 333 FPSVG----VFNPGGAIWIADEHPRPYRTLS-----DVPF---VRGGFGNPH 373

Qy 426 N-----TPERTANYQSPYESPGIQLKDSSETLPEPETERPNYESYSHRLSHI----- 473

Db 374 YVLGLRGVAFQQTGHTRTFRNSG--TIDSLDEIPQDMSGAPWNDYSHLVNHTVFRV 431

Qy 474 -GLISQS-RVHVSVYVWTHRSADRTNITSSDSTOIPLVKSFNLNGSTSVVSGPFGTGGD 531

Db 432 PGELSGSDSWRAWPFWSWTHRSATPTNIDPERITQIPLVKAHTLQSGTIVVRGPGTGGD 491

Qy 532 IITNVNGSVLSLGLFNNTSLQRYRVRYAASQTMVLRVTVGSGSTTFPQGFPTMSAN 591

Db 492 ILARTSGPPAYITVINGQLPQRYARIRYASTTNLRIVYTVAGERIFAGQFNKMTDTC 551

Qy 592 ESLSQSFRPAEPVGVISAGSGTAGISISNNAGROTFF-----HFDKIEFIPITATFEA 645

Db 552 DPLTFQSFSYATINTAFTFPMSSQ-----SFTVGADTFSSGNEVYIDRFELIPVATFEA 606

Qy 646 EYDLERAQEAVALFTNTNPRRLKTDVTDYHIDQVSNLVACLDFCLDEKRELEKVKY 705

Db 607 EYDLERAQKAVNALFTSIQIGTKTDVTDYHIDQVSNLVACLDFCLDEKRELEKVKY 666

Qy 706 AKLSDERNLLODPNFTSINKQPDFISTNEQSNFTSIHQSEHGWSGSENIITIQEGDNVF 765

Db 667 AKLSDERNLLODPNFKGINRQLD-----RGMWGSDTITIQGDDVF 708

Qy 766 KENYVTLPGTFNECYPTLYOKIGESKAYTRYQYLRGYIEDSDLEIYLIRYNAKHETL 825

Db 709 KENYVTLPGTFDECYPTLYOKIDESKAYTRYQYLRGYIEDSDLEIYLIRYNAKHETV 768

Qy 826 DVPGTESVWPLSVESPIGRGPNRCAPHEFNWPDLDSCRDGKCAHSHHFSLDIDIG 885

Db 769 NVLGTCSLWPLSVQSPIRKGPENRCAPHEFNWPDLDSCRDGKCAHSHHFSLDIDVG 828

Qy 886 CTDLHENLGVVWVFKIKTQEGHARLGNLFIEBKPLGLBALSRVRAEKAKKWDKREKLOL 945

Db 829 CTDLNELDVWVIFKIKTQEGHARLGNLFIEBKPLGLBALSRVRAEKAKKWDKREKLEL 888

Qy 946 ETKRVYTEAKEADALFVDSQYRLQADTNIGMHAADKLVRHREAYLSLSLPGVNA 1005

Db 889 ETNIVYKEAKESVDALFVNSQYDQLQADTNIGMHAADKRVHRIREAYLPELSVIPGVNV 948

Qy 1006 EIPFEELEGRIITAFSLYDARNVYKRGDFNNGLACVNVKGVHDV-QQSHRSVLVPIPEWA 1064

Db 949 DIFEELKGRIFTAFPLYDARNVYKRGDFNNGLACVNVKGVHDVVEEQNNHRSVLVPEWA 1008

Qy 1065 EVSQAVRVCPRGYILRVYAYKEGEGCVTHIEINNTDELKFKNCBEEVYPTDGTG 1124

Db 1009 EVSQEVRVCPRGYILRVYAYKEGEGCVTHIEINNTDELKFKNCBEEVYPTDGTG 1068

Qy 1125 NDYTAHQ---GTAVCNRSNAGYEDAYEDVTTASVNVKPYEETVTDVRRDNHCEYDRGY 1181

Db 1069 NDYTAHQEYGGAY-TSRNGVDETYGNSVVPADYASVYEKSYTDGRRDPCESNRGY 1127

Qy 1182 VNPPLPAGYMTKELEYFPETDKVWIEIGETGKFIIVDSVELLMEE 1228

Db 1128 GDYTPLPAGYMTKELEYFPETDKVWIEIGETGKFIIVDSVELLMEE 1174

RESULT 11

S32647

parasporal crystal protein cryIa1 - Bacillus thuringiensis

C;Species: Bacillus thuringiensis

C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004

C;Accession: S32647

R;Lambert, B.

submitted to the EMBL Data Library, April 1993

A;Reference number: S32645

A;Accession: S32647

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1160 <Lam>

A;Cross-references: UNIPROT:Q45747; EMBL:Z22511; NID:g295963; PIDN:CAA80234.1; PID:g2959

C;Superfamily: parasporal crystal protein

C;Keywords: Delta-endotoxin

Query Match 52.6%; Score 3411; DB 2; Length 1160;

Best Local Similarity 56.2%; Pred. No. 6.3e-209; Indels 108; Gaps 21;

Matches 695; Conservative 145; Mismatches 288; Indels 108; Gaps 21;

Qy 22 NHSTQMDLSPDARIEDSLCIAEGNNINP---LVASASTVOTGINTAGRIGLVGLG-----V 72
Db 4 NHQNO-----CIPYNCLSNPDAILLDAERLETGTVADISLGLINFLYSNFV 50
Qy 73 PFAGQIASFYSFLVGLWPRGRDOWEIFLEHVEQLINQOITENARNTALARLOGLSFR 132
Db 51 PGGGFIVGLLELWFGVP---SQWEIFLAQIEQLISQRIEEFARNQAISREGLSNYE 107
Qy 133 AYQOSLEDWLENRRDARTSRVLYTQYIALELDLFLNAMPFAIRNOEVPFLMVYAAQANLH 192
Db 108 IYETFPANEKDPSPALREEMRTQFNWMSALIAAPLLLRVNYEVALLSVYVQAANLH 167
Qy 193 LLLRDASLFGSEFGLTSQBIQRYRQVEQTRDYSYCVYEWNTGLNLRGTNAASVVR 252
Db 168 LSVLRDVSVYQORGPDPATVNSRYSDLTRLIHVYTDHCVDTYNDGLKLEGRSLDWWV 227
Qy 253 YNQRRDLTLGLDLVALFSDYDTRTPINTSAQLTREVYTD---ICATGVNMAAMWYN 310
Db 228 YNFRRLTLTSLVDIIAFFFNIDYEAPYIQTASQLTREVLDLFPVNETLSPASV--- 283
Qy 311 NNAPSFSAIETAVIRSHLLDLFQLTIFSSRWASATRMVYRGHTIOSRPTGGGLN- 369
Db 284 ---PTFSAAESAILRSHLDVFLNFTIYDS-----LASVAYMGHGLVNSFRGTGTTNL 335
Qy 370 --TSTHGSTNTSINPVRLSFF-SDVYWTBYSYAGVLLWGI-YLEPIHGVTVPFRNPNQ 425
Db 336 IRSPLYREGNTERPVTISAPSPVPIFRTLISY-----FTGLNNNPVAGIEGVF- ---Q 386
Qy 426 NTERGTANTSQPVESQGLQKQSETELPPETTERPNYESYSHRLSH---IGLISQSRVH 482
Db 387 NTISRSIYRKSGPT-----DSFSELPPQDVSVSPAIGYSHRLCHATEFLERISGPRIA 438
Qy 483 VPVYSWTHRSADRNTTSSDITQIPLVKSFNLSGTSVWSGPGTGGDIIRTNVNSVL 542
Db 439 GTVFSWTHRSASPINEVSPSRITQIPWKAKHTLASGASVINGPGFTGGDILTRNSMGDLG 498
Qy 543 SMGLNFNNTSLQRYRVRVRYA---ASQTMVLRVTVGGSTTFDQGPPTWSANESLTSQSR 600
Db 499 ALRVTFGRLPQSYIIFRVASVANRSGTFRYSQPPS---YGISFPKTMAGEALTSRSPA 556
Qy 601 FAEPFVGISAGSQ-----TAGISINNAGROTFFHDKIEFTIPTATFEAEVDLERAQ 653
Db 557 HTTLFTPTFSRAQEEFDLYTQSGV-----YIDRIEFIPVDATFESEINLERAQ 605
Qy 654 EAVNALFTNTPRELKTDVTDYHIDQVSNLVACLSDEFCLDEKKELEKVKYAKRLSDER 713
Db 606 KAVNALFTSNQLGLKTDVTDYHIDQVSNLVACLSDEFCLDEKKELEKVKYAKRLSDER 665
Qy 714 NLLQDPNFTSINKOPDFISTNEQSNFTSIHQSHGHWGSENIITIQBGNDVFKENYVTLF 773
Db 666 NLLQDPNFRGNRQPD-----RGRWGSTDTITIQQGDVDFKENVYTLT 707
Qy 774 GTFNECYPTLYQKIGSELKAYTRYQLRGVIEDSQDLEIYLIRYNAKHETLDVPGTESV 833
Db 708 GTFDECYPTLYQKIDESKLKAYTRYQLRGVIEDSQDLEIYLIRYNAKHETLVNVPGTGSL 767
Qy 834 WPLSVESPIGRCGPENRCAPHEFNWPDLCSCRGEKCAHSHHFSLDIDIGCTDLHENL 893
Db 768 WPLSVQSPIGKCGEPNRCAPHEFNWPDLCSCRGEKCAHSHHFSLDIDVGCDDLNEEDL 827

Qy 894 GVMVVFVKIKTOEGHARLGNLEPIEEKPLLGREALSRVRAKSKWRDKREKLOLETKRVYTE 953
Db 828 GVMVIFKIKTQDGHARLGNLEPIEEKPLVGEALARVRAKSKWRDKREKLELETNIVYKE 887
Qy 954 AKCAVDALFVDSQVNRQADTNLGMIAADKLVHRIREAVLSLSVLPVGNAAEIPFEELEG 1013
Db 888 AKESVDALFVNSQYDQLOADTNLAMIHAADKRVHSIREALYLPUSLVIPGVNAGIPFEELEG 947
Qy 1014 RIITAIISLYDARNVVKNGDFNNGGLACWNVKGHDV- QQSHRSRVLVIPWEAEYSQAVRV 1072
Db 948 RIFATSLYDARNVVKNGDFNNGGLSCWNVKGHDVDSQNNHRSVLVPEWEAEYSQAVRV 1007
Qy 1073 CPGRYILRVYAKYEGYEGCVTIHEIENNTDELKFNKCEEEVYPTDTGTGTCNDYTAHQ 1132
Db 1008 CPGRYILRVYAKYEGYEGCVTIHEVDNNTDELKFSNCEKEQYVPGNTVACNDYNKNHG 1067
Qy 1133 TAVCNSNAGVEADAYEVDTTASVNYKPYEETTVDRRDNHCEYDRGVNYPPLPAGYM 1192
Db 1068 ANACSSNRNGYDESYESNSSIPADYAPVYBEEAYTDGQRNCPCEFNRGHT---PLPAGYV 1124
Qy 1193 TKELEYFPETDKVMIEIGETEGKPIVDSVELLLMEE 1228
Db 1125 TAELEYFPETDVTWVEIGETEGPIVDSVELLLMEE 1160

RESULT 12

JC2219

parasporal crystal protein cryIIa - Bacillus thuringiensis

C;Species: Bacillus thuringiensis

C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004

C;Accession: JC2219

R;Udayasuriyan, V.; Nakamura, A.; Mori, H.; Masaaki, H.; Uozumi, T.

Biosci. Biotechnol. Biochem. 58, 830-835, 1994

A;Title: Cloning of a new cryIIA(a) gene from Bacillus thuringiensis strain FU-2-7 and an

A;Reference number: JC2219; MUID:94289859; PMID:7764972

A;Accession: JC2219

A;Molecule type: DNA

A;Residues: 1-1176 <DA>

A;Cross-references: UNIPROT:Q45736; DDBJ:D17518; NID:G506190; PIDN:BAA04468.1; PID:G53571

C;Genetics:

A;Gene: cryIIA(a)

C;Superfamily: parasporal crystal protein

C;Keywords: Delta-endotoxin

Query Match 52.6%; Score 3405.5; DB 2; Length 1176;

Best Local Similarity 56.3%; Pred. No. 1.5e-208;

Matches 697; Conservative 149; Mismatches 293; Indels 100; Gaps 24;

Qy 27 MDLSDPARIBDSLCAEGNNINP---LVASASTVOTG---INIAGRILGVL---GVFPAGQ 77

Db 1 MDNNPNI---NECIPYNCLSNPEVEVLGGERIETGVTPIDLSLTQFLLSFVPAGAG- 55

Qy 78 IASFYSFLVGLWLP-RGRDOWEIFLEHVEQLINQOITENARNTALARLOGLSDFRAYQQ 136

Db 56 ---FVLGLVDIINGIFGFSQWDAFLVQIQIEQINRIIEEFARNQAISREGLSNLYQIYAE 112

Qy 137 SLEDWLENRRDARTSRVLYTQYIALELDLFLNAMPFAIRNOEVPFLMVYAAQANLHLLL 196

Db 113 SFREWEADPTNPALREEMRIQFNDMNSALITATPLLAVQNVQVFLSVYVQAANLHLSVL 172

Qy 197 RDASLFGSEFGLTSQBIQRYRQVEQTRDYSYCVYEWNTGLNLRGTNAASVRYNQF 256

Db 173 RDVSFQGRWGFDAATINSRYNDLTRIGNVTDYAVRWYNTGLERVNGPDSRDWRVYNQF 232

Qy 257 RRDITGLVDLVALFSDYDTRTPINTSAQLTREVYTDATCATGVNMAAMWYNNAAPSF 316

Db 233 RRELTFLVDLVALFSDYDTRTPINTSAQLTREVYTDATCATGVNMAAMWYNNAAPSF 280

Qy 317 SAETAV---IRSPHLLDFLEQLTIFSTSRKWSATRMVYRGHTIOSRPIGCG---LNT 370

Db 281 RGMARIEQNIQPHLMDILNSITTYT-----DVHRGFNYWSHQITASPVFGSPGFAP 335

A;Title: Transcriptional and translational start sites for the Bacillus thuringiensis cp
 A;Reference number: A92410; MUID:83109004; PMID:6296116
 A;Accession: A03489
 A;Molecule type: DNA
 A;Residues: 1-76, 'P', 78-147, 'L', 149-247, 'S', 249-282, 'M', 284-285, 'R', 287-288, 'QN', 291-292
 A;Experimental source: strain HD-1
 C;Comment: This protein is present in crystalline form as a component of the spore coat.
 C;Superfamily: parasporal crystal protein
 C;Keywords: delta-endotoxin

Query Match 52.5%; Score 3399.5; DB 1; Length 1178;
 Best Local Similarity 56.7%; Pred. No. 3.5e-208;
 Matches 704; Conservative 148; Mismatches 287; Indels 102; Gaps 26;

Qy	27	MDLSPDARIEDSLCAEINNINP---LVSASTVQTG---INIAGRILGVL---GVPEAGQ 77
Db	1	MDNNFNI---NECIPYCNLSNPEVEVLGGERTGTPTDIDLSLTFQLLSBFVPAG- 55
Qy	78	IAGFYSFLVGLWLP-RGRDQWEIPLFHEVQQLINQIITENARNTALARLOGLGDSFRAYQQ 136
Db	56	---FVLGLVDIIWGIQFQSDWDAFLVQIEQLINQRIEFAFNQALISEGLSNLYQIYAE 112
Qy	137	SLDWLENRDDARTSRVLYTQYIALEDPLNAPLFAIRNOEVPPLMVYAAANLHLLLL 196
Db	113	SFEWEADPTNPALREEMRIQFNDMSALTATPLFAVQYQVPLLSVYVQAANLHLSVL 172
Qy	197	RDASLFGSEFGLTSQEIQRVYRQVEQTRDYSVCVEMWNTGLNSLGRGTNAASWVRYNQF 256
Db	173	RDVSFQORGFDAATNSRYNDLTRLIGNYTDYAVRWYNTGLERWGPDSRDWRVYNQF 232
Qy	257	RBDLTGLVDLVALFPYSYDTRTPINTSAQLTREYVTDALCATGVNMAAMNNYNNAPSF 316
Db	233	RRELTTLVDLVALFPNYDSRRYPRTVSQLTREIYT-----NPVLENFDGSPGSA 284
Qy	317	SAIETAVIRSHLLDFLEQLTIFSTSRWSATHMTWRGHTIQSRPIGG---LNTSTH 373
Db	285	QGIERS-IRSPHMLDIILNSITIVT-----DAHRGYYWSGHOIMASPVGSGPEFTPLY 338
Qy	374	GSTNTSINPVL-SFFSRDVTWTESYAGVLLGYLEPIHGVPTRVFNPNPO-----N 426
Db	339	GTWGNAAPORIVAQLCGGYRTLST-----LYRPFTN-----IGINNQLSVLDGT 386
Qy	427	TFERGTYNY--SOPYESFGLKDSFETLPPETERENYESYSHRLSHIGLI-----S 477
Db	387	EFAYGTSNLPSSAVYRKSG--TVDSLDEIIPPNNVPPROGFSHRLSHVSMFSGFSNS 444
Qy	478	QSRVHVPIVSWTHRSADRTNTISDSITQPLVKSNLNSGTSVSGPGFTGGDIIRTNV 537
Db	445	VSIRAPMFWSIHRSAFNNIIASDSITQIPAVKGNFLNG-SVISGPGFTGGDLVRLNS 503
Qy	538	NGS-----VLSMGLNFNNTSLORYRVRYAASQTMVLAVTVGGSTPDQGFPTMSAN 591
Db	504	SGNNIQNRGYIEVPIHPPTS--TRYRVRYASVTPHILNVNNGNSIFNTVPTATSL 562
Qy	592	ESLTSQSFEAFEPFVGISASGSGTAGISINNAGROTFHFDEKIEFIPITATFEAEYDLR 651
Db	563	DNLQSSDFGYFESANAFSTSLGNIVG--VRNFSGTAGVIIDREFIPVATLEAYNLER 620
Qy	652	AQAVNALFTNTNPRILKTDVTHIDQVSNLVACLSDPECLDEKRELLKVKYAKRLSD 711
Db	621	AQAVNALFTSNQLGKTNVTDVTHIDQVSNLVYLSDEFCLDEKRELSKVKHAKRLSD 680
Qy	712	ERNLQDPNFTSINKOPDFISTNEQSNFTSIHQSESHGWMGSENITIOEGNDVFKENYVT 771
Db	681	ERNLLQDSNFKDINRQ-----ERWGGSGTITIQGGDDVFKENYVT 722
Qy	772	LPGTNECYPTLYQKIGESLKYTRYQLRGYIEDSQDLEIYLIRYNAKHETLDVPGTE 831
Db	723	LSGTFDECYPTLYQKIDESKLAFTRYQLRGYIEDSQDLEIYLIRYNAKHETVNVPGTG 782
Qy	832	SVNPLSVESPIGRGEPNRCAPHFNWPDLDCCSRGCKCAHSHHFSLDIDIGCTDLHE 891
Db	783	SLWPLSAQSPIGKGEPNRCAPHLNPNPDLDCCSRGCKCAHSHHFSLDIDVGCITDLNE 842

Search completed: April 21, 2005, 04:17:21
 Job time : 59 secs

Qy	892	NLGVWVFKIKTOEGHARLGNLEFIEKPLLGALSRVKBAEKKWRDKREKLOLETKRVY 951
Db	843	DLGVWVFKIKTOEGHARLGNLEFIEKPLVUGALARKKWRDKREKLEWEINIVY 902
Qy	952	TEAKEAVDALFVDSQYNRLQADTNIGMHAADKLVHRIEAYLSLSVIPGVNAEIFEEL 1011
Db	903	KEAKESVDALFVNSQYDQLQADTNIAIHAADKRVHSIREAYLPESLVIPGVNAEIFEEL 962
Qy	1012	EGRIITAIISLYDARNVVKNGDFNGLACWNVKGHDV-QQSHHRSVLVIPDEWAEVSQAV 1070
Db	963	EGRIITAFSLYDARNVIKNGDFNGLSCWNVKGHDVVEQNNQSRVSVLVPEWAEVSQEV 1022
Qy	1071	RVCPRGRGYTLRVYAKGEGGCVTTHEENNTDELKFKNCEEEVYPTDTGTCNDYTAH 1130
Db	1023	RVCPRGRGYTLRVYAKGEGGCVTTHEENNTDELKFSNCVEEYYPNNVTTCNDYTVN 1082
Qy	1131	Q---GTAVCNRSNAGYEDAYEDVDTTASVNYKPTYEBETTDVRRDNHCEYDRGYVNYPPL 1187
Db	1083	QBEYGGAY-TSRNRGYNEAPSV---PADYASVYEKSYTDGRRENPCFENRGYDYTPL 1137
Qy	1188	PAGYMTKELEYPPETDKWIEIGETEGKFIIVDSVELLLMEE 1228
Db	1138	PVGVTKELEYPPETDKWIEIGETEGTETGFIIVDSVELLLMEE 1178

This Page Blank (uspto)

QY 541 VLSMGLNFNTSLQRYRVRVRYAASQTMVLRTVTVGGSTTDFDQGPSTMSANESLTSQSPR 600
DB 541 VLSMGLNFNTSLQRYRVRVRYAASQTMVLRTVTVGGSTTDFDQGPSTMSANESLTSQSPR 600
QY 601 FAEFPVGISASGQTAGISISNAGROTFFDKIEFIPITATPEAYDLERAOEAVNALF 660
DB 601 FAEFPVGISASGQTAGISISNAGROTFFDKIEFIPITATPEAYDLERAOEAVNALF 660
QY 661 TTNTPRLKTDVTDYHDQVSNLVACLSDFCLDEKRELEKVKYAKRLSDERNLLQDPN 720
DB 661 TTNTPRLKTDVTDYHDQVSNLVACLSDFCLDEKRELEKVKYAKRLSDERNLLQDPN 720
QY 721 FTSINQKPDISTNEOSNFTSIHQSHGHWGSENITIQEGNDVFKENYVTLPGTFNECY 780
DB 721 FTSINQKPDISTNEOSNFTSIHQSHGHWGSENITIQEGNDVFKENYVTLPGTFNECY 780
QY 781 PTLYQKIGESSELKAYTRYQLRGVIEDSQDLEIYLIRYNAKHETLDVPGTESVWPLSVES 840
DB 781 PTLYQKIGESSELKAYTRYQLRGVIEDSQDLEIYLIRYNAKHETLDVPGTESVWPLSVES 840
QY 841 PIGRCGEPNRCAPHFENWPDLDSCSRDGEKCAHSHHFSLDIDIGCTDLHENLGVWVVK 900
DB 841 PIGRCGEPNRCAPHFENWPDLDSCSRDGEKCAHSHHFSLDIDIGCTDLHENLGVWVVK 900
QY 901 IKTOEGHARLGNLEFTEEPKLLGEALSRVRAEKKWRDKREKLOLETYKRYVYTAKEAVDA 960
DB 901 IKTOEGHARLGNLEFTEEPKLLGEALSRVRAEKKWRDKREKLOLETYKRYVYTAKEAVDA 960
QY 961 LFDVDSQVNRLOADTNIGMIHAADKLVHRIEAYLSELSVTPGVNAEIFELEGRIITAIS 1020
DB 961 LFDVDSQVNRLOADTNIGMIHAADKLVHRIEAYLSELSVTPGVNAEIFELEGRIITAIS 1020
QY 1021 LYDARNVKNXGDFNGLACVNVKXGHVDVQOHSRSLVLPWEABRVSAQVRCVPCRGYIL 1080
DB 1021 LYDARNVKNXGDFNGLACVNVKXGHVDVQOHSRSLVLPWEABRVSAQVRCVPCRGYIL 1080
QY 1081 RVYAYKEGEGCVTTHEINNTDELKFNCEBEEVPTDTGTCNDYTAHQGTAVCNERN 1140
DB 1081 RVYAYKEGEGCVTTHEINNTDELKFNCEBEEVPTDTGTCNDYTAHQGTAVCNERN 1140
QY 1141 AGVEDAYEVDVTTASVNVKPYEBEETVDRDNHCEYDRGVNYPPLPAGYVTKLEYFP 1200
DB 1141 AGVEDAYEVDVTTASVNVKPYEBEETVDRDNHCEYDRGVNYPPLPAGYVTKLEYFP 1200
QY 1201 ETDKWIEIGETEGTFIVDSVLELLMEE 1228
DB 1201 ETDKWIEIGETEGTFIVDSVLELLMEE 1228

RESULT 3
US-10-428-961-63
; Sequence 63, Application US/10428961
; Publication No. US2003023711A1
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Ruper, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin
; TITLE OF INVENTION: Polynucleotides, Compositions, and Methods of Use (Amended)
; FILE REFERENCE: MEC0201-1
; CURRENT APPLICATION NUMBER: US/10/428,961
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/661,322
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: 60/153,995
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 63
; LENGTH: 1227

; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
; US-10-428-961-63

Query Match 91.3%; Score 5912.5; DB 15; Length 1227;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1127; Conservative 36; Mismatches 62; Indels 5; Gaps 3;

QY 1 LTSNRKNEEIIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
DB 1 LTSNRKNEEIIINALSIPAVSNHSAQMNLSLDARIEDSLCIAEGNNIDPPVSASTVQTGI 60
QY 61 NIAGRILGVLPVFPAGQIASFYSLVGLMELPRGRDQWEIFLEHVQLINOQIITENARNTA 120
DB 61 NIAGRILGVLPVFPAGQIASFYSLVGLMELPRGRDQWEIFLEHVHLIRQOVTENTRDTA 120
QY 121 LARLOGLGDSFRAYQOSLEDWLENRDDARTSRVLYTQYIALELDFLNAMPLFAIRNOEVP 180
DB 121 LARLOGLGDSFRAYQOSLEDWLENRDDARTSRVLYTQYIALELDFLNAMPLFAIRNOEVP 180
QY 181 LLMVYAAANLHLLLRDASLFGSEFGLTSQEIQRYRQVEQTRDYSYCVWEYNTGLN 240
DB 181 LLMVYAAANLHLLLRDASLFGSEFGLTSQEIQRYRQVEQTRDYSYCVWEYNTGLN 240
QY 241 SLRGTVAAWRYNQPRDLITGLVDLVALFPSYDTRTPTINTSAQLTREYVYDAIGATG 300
DB 241 NLRGTVNAESWRYNQPRDLITGLVDLVALFPSYDTRTPTINTSAQLTREYVYDAIGATG 300
QY 301 V--NMAWMVNNAPSFSAIETAVIRSPLLDFLQLTIFSTSSSWSATRHMTYWRGHT 358
DB 301 APSGFATWNNAPSFSAIETAVIRSPLLDFLQLTIFSTSSSWSATRHMTYWRGHT 358
QY 359 IQSRPTGGGLNTSTHGSTNTSINPVLSPFSDRVYVTVSGYVLLMGVILEPIHGVPTVR 418
DB 361 LESRTIRGSLTWTGNTSINPVLSPFSDRVYVTVSGYVLLMGVILEPIHGVPTVR 418
QY 419 FNRFPQNTFEGTANYSQPYESPLQKDSQETELPPTETERNPYESYSHRSLHIGLISQ 478
DB 419 FNRWNPNSL-RGSLLYTTIGYGVGTQLPFDSQETELPPTETERNPYESYSHRSLHIGLISQ 477
QY 479 SRVHVPTVSWTHRSADRTNTISSDSITQIPVKSFNLSGTSVWSGPGTGGDIITNVN 538
DB 478 NTLRAPVSWTHRSADRTNTISSDSITQIPVKSFNLSGTSVWSGPGTGGDIITNVN 537
QY 539 GSVLWGLNFNTSLQRYRVRVRYAASQTMVLRTVTVGGSTTDFDQGPSTMSANESLTSQS 598
DB 538 GSVLWGLNFNTSLQRYRVRVRYAASQTMVLRTVTVGGSTTDFDQGPSTMSANESLTSQS 597
QY 599 FRPAEFPVGISASGQTAGISISNAGROTFFDKIEFIPITATPEAYDLERAOEAVNA 658
DB 598 FRPAEFPVGISASGQTAGISISNAGROTFFDKIEFIPITATPEAYDLERAOEAVNA 657
QY 659 LFTNTNPRRLKTDVTDYHDQVSNLVACLSDFCLDEKRELEKVKYAKRLSDERNLLQD 718
DB 658 LFTNTNPRRLKTDVTDYHDQVSNLVACLSDFCLDEKRELEKVKYAKRLSDERNLLQD 717
QY 719 PNFTSINKQPDISTNEOSNFTSIHQSHGHWGSENITIQEGNDVFKENYVTLPGTFNE 778
DB 718 PNFTSINKQPDISTNEOSNFTSIHQSHGHWGSENITIQEGNDVFKENYVTLPGTFNE 777
QY 779 CYPTLYQKIGESSELKAYTRYQLRGVIEDSQDLEIYLIRYNAKHETLDVPGTESVWPLSV 838
DB 778 CYPTLYQKIGESSELKAYTRYQLRGVIEDSQDLEIYLIRYNAKHETLDVPGTESVWPLSV 837
QY 839 ESPICRCGEPNRCAPHFENWPDLDSCSRDGEKCAHSHHFSLDIDIGCTDLHENLGVWV 898
DB 838 ESPICRCGEPNRCAPHFENWPDLDSCSRDGEKCAHSHHFSLDIDIGCTDLHENLGVWV 897
QY 899 FKIKTOEGHARLGNLEFTEEPKLLGEALSRVRAEKKWRDKREKLOLETYKRYVYTAKEAV 958
DB 898 FKIKTOEGHARLGNLEFTEEPKLLGEALSRVRAEKKWRDKREKLOLETYKRYVYTAKEAV 957
QY 959 DALFVDSQVNRLOADTNIGMIHAADKLVHRIEAYLSELSVTPGVNAEIFELEGRIITA 1018

This Page Blank (uspto)

Db 2161 TTCATCCATCAATAAGCAACAGACTTCATATCTACTAATGAGCAATCGAATTTTCA 2220
 Qy 2221 TCATATCATGAACAATCTGAACATGGATGGTGGGAGTGCAGACATTAATCCAGGAA 2280
 Db 2221 TCATATCATGAACAATCTGAACATGGATGGTGGGAGTGCAGACATTAATCCAGGAA 2280
 Qy 2281 GGAATGACGTATTTAAAGAGAAATTAAGTGCATACCTACCGGGGACTTTTAATGAGTGTAT 2340
 Db 2281 GGAATGACGTATTTAAAGAGAAATTAAGTGCATACCTACCGGGGACTTTTAATGAGTGTAT 2340
 Qy 2341 CCGAGGTATTTAATCAAAAAATAGAGAGTGCAGAAATTAAGTGCATACCTACCGGGGACTTTTAATGAGTGTAT 2400
 Db 2341 CCGAGGTATTTAATCAAAAAATAGAGAGTGCAGAAATTAAGTGCATACCTACCGGGGACTTTTAATGAGTGTAT 2400
 Qy 2401 TTAAGGGGTATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAATGCG 2460
 Db 2401 TTAAGGGGTATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAATGCG 2460
 Qy 2461 AAACATGAAACATTTGATGTTCCAGGTACCGAGTCCGATGGCGGCTTTTCAGTTGAAAGC 2520
 Db 2461 AAACATGAAACATTTGATGTTCCAGGTACCGAGTCCGATGGCGGCTTTTCAGTTGAAAGC 2520
 Qy 2521 CCAATCGAAGGTGCGAGAACCGAATCGATGCGCACCAATTTTGAATGGAATCCTGAT 2580
 Db 2521 CCAATCGAAGGTGCGAGAACCGAATCGATGCGCACCAATTTTGAATGGAATCCTGAT 2580
 Qy 2581 CTAGATGTTCTGCGAGAGTGCAGAAATGCGCATCAATTTCCATCATTTCTCTTTG 2640
 Db 2581 CTAGATGTTCTGCGAGAGTGCAGAAATGCGCATCAATTTCCATCATTTCTCTTTG 2640
 Qy 2641 GATATTGATATTGGATGTCACAGACTTCATGAGAACTAGGCGGTGGTGGTATTCAAG 2700
 Db 2641 GATATTGATATTGGATGTCACAGACTTCATGAGAACTAGGCGGTGGTGGTATTCAAG 2700
 Qy 2701 ATTAAGACGAGGAGGTTCATGCAAGACTAGGAACTCTGGAATTTATGAAGAGAAACCA 2760
 Db 2701 ATTAAGACGAGGAGGTTCATGCAAGACTAGGAACTCTGGAATTTATGAAGAGAAACCA 2760
 Qy 2761 TTATTAGGAGACACTGCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
 Db 2761 TTATTAGGAGACACTGCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
 Qy 2821 GAAAACTCAATTTGAAACAAACAGATATATACAGAGGCGCAAGAGAGAGAGAGAGAGAG 2880
 Db 2821 GAAAACTCAATTTGAAACAAACAGATATATACAGAGGCGCAAGAGAGAGAGAGAGAGAG 2880
 Qy 2881 TTATTGAGATTTCTCAATATATAGATTACAGCGGATACAAACATTTGGCATGATTCAT 2940
 Db 2881 TTATTGAGATTTCTCAATATATAGATTACAGCGGATACAAACATTTGGCATGATTCAT 2940
 Qy 2941 GCGGAGATAACTGTTGATTCGAAATTCGAGAGGCTTATCTGTCAGAAATTTCTGTTATC 3000
 Db 2941 GCGGAGATAACTGTTGATTCGAAATTCGAGAGGCTTATCTGTCAGAAATTTCTGTTATC 3000
 Qy 3001 CCGGGTAAATCGGAAATTTTGAAGATTAAGAGTGCATATCACTGCAATCTCC 3060
 Db 3001 CCGGGTAAATCGGAAATTTTGAAGATTAAGAGTGCATATCACTGCAATCTCC 3060
 Qy 3061 CTATACGATCGGAGAAATGCTGTTAAATGTTGATTTTAAATGAGTATGAGTATGCTG 3120
 Db 3061 CTATACGATCGGAGAAATGCTGTTAAATGTTGATTTTAAATGAGTATGAGTATGCTG 3120
 Qy 3121 AATGTAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
 Db 3121 AATGTAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
 Qy 3181 GATGGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
 Db 3181 GATGGAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
 Qy 3241 CCGTGTACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
 Db 3241 CCGTGTACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300

Qy 3301 AACAATACAGAGCACTAAATTTTAAATCTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 3360
 Db 3301 AACAATACAGAGCACTAAATTTTAAATCTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 3360
 Qy 3361 ACAGGAACGTGTAAATGATATATCTGCACCAAGGTACAGCAGTATGTAATTTCCCGTAAT 3420
 Db 3361 ACAGGAACGTGTAAATGATATATCTGCACCAAGGTACAGCAGTATGTAATTTCCCGTAAT 3420
 Qy 3421 GCTGGATATGAGATGATATGATATGATATGATATGATATGATATGATATGATATGATAT 3480
 Db 3421 GCTGGATATGAGATGATATGATATGATATGATATGATATGATATGATATGATATGATAT 3480
 Qy 3481 TATGAAG 3540
 Db 3481 TATGAAG 3540
 Qy 3541 TATGTGAATTTATCCACCACTACAGCTGCTTATATGACAAAAGAAATTAGAATATCTTCCCA 3600
 Db 3541 TATGTGAATTTATCCACCACTACAGCTGCTTATATGACAAAAGAAATTAGAATATCTTCCCA 3600
 Qy 3601 GAAACCGATAGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 3660
 Db 3601 GAAACCGATAGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 3660
 Qy 3661 GTGGATTTACTCTTATGAGAGAGATAG 3687
 Db 3661 GTGGATTTACTCTTATGAGAGAGATAG 3687

RESULT 2
 US-10-428-961-37
 ; Sequence 37, Application US/10428961
 ; Publication NO. US20030237111A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baum, James A.
 ; APPLICANT: Chu, Chih-Rei
 ; APPLICANT: Donovan, William P.
 ; APPLICANT: Gilmer, Amy J.
 ; APPLICANT: Rupar, Mark J.
 ; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin
 ; TITLE OF INVENTION: Polynucleotides, Compositions, and Methods of Use (Amended)
 ; FILE REFERENCE: MECO201--1
 ; CURRENT APPLICATION NUMBER: US/10/428,961
 ; PRIOR FILING DATE: 2003-05-02
 ; PRIOR APPLICATION NUMBER: 09/661,322
 ; PRIOR FILING DATE: 2000-09-13
 ; PRIOR APPLICATION NUMBER: 60/153,995
 ; PRIOR FILING DATE: 1999-09-15
 ; NUMBER OF SEQ ID NOS: 63
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 37
 ; LENGTH: 4173
 ; TYPE: DNA
 ; ORGANISM: Bacillus thuringiensis
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1) ... (3687)
 US-10-428-961-37

Query Match 99.4%; Score 3666.2; DB 17; Length 4173;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 3674; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 Qy 1 TTGACTTCAATAGCAAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 60
 Db 1 TTGACTTCAATAGCAAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 60
 Qy 61 TCGAATCAATTCACACAAATGAGATCTATCACCAGATGCTCGTATTCAGAGATCTTTTGTGT 120
 Db 61 TCGAATCAATTCACACAAATGAGATCTATCACCAGATGCTCGTATTCAGAGATCTTTTGTGT 120
 Qy 121 ATAGCCGAGGGGAGATTAATCAATCACTTGTAGGAGATCAACAGTCCAAACCGGATATT 180

This Page Blank (uspto)

121 ATAGCCGAGGGAATATATCAATCACTTGTAGGCAATCAAGTCCAAACGGGTATT 180
181 AACATAGCTGGTAGAATACTAGGTGTAATAGCGGTACCGTTTGGTGGCAAAATAGCTAGT 240
181 AACATAGCTGGTAGAATACTAGGTGTAATAGCGGTACCGTTTGGTGGCAAAATAGCTAGT 240
241 TTTTATAGTTTCTTGTGGTGAATATAGCGCCCGCGGCGAGAGTCACTGGGAAATTTTC 300
241 TTTTATAGTTTCTTGTGGTGAATATAGCGCCCGCGGCGAGAGTCACTGGGAAATTTTC 300
301 CTAGAACATGTCGAACAACCTTAAATCAACAATAACAGAAAATGCTAGGAATACGGCA 360
301 CTAGAACATGTCGAACAACCTTAAATCAACAATAACAGAAAATGCTAGGAATACGGCA 360
361 CTTGCTCGAATACAAAGTTTAGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT 420
361 CTTGCTCGAATACAAAGTTTAGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT 420
421 TGGCTAGAAAACCGTGATGATCAAGAACGAGAGTGTCTTTATACCCCAATATATAGCC 480
421 TGGCTAGAAAACCGTGATGATCAAGAACGAGAGTGTCTTTATACCCCAATATATAGCC 480
481 TTAGAACTTTGATTTCTTAATCGATGCGCTTTTTCGCAATTAGAAAACCAAGAGTTCCA 540
481 TTAGAACTTTGATTTCTTAATCGATGCGCTTTTTCGCAATTAGAAAACCAAGAGTTCCA 540
541 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACTTATATATAGAGATGCTCT 600
541 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACTTATATATAGAGATGCTCT 600
601 CTTTGTGTAGTAATTTGGCTTACATCGCAGGAATTCAGCTTATATGAGCGCCAA 660
601 CTTTGTGTAGTAATTTGGCTTACATCGCAGGAATTCAGCTTATATGAGCGCCAA 660
661 GTGGAAACAACGAGAGATTTCCGACTATGCTAGATGCTGATATATACAGGTCTTAAT 720
661 GTGGAAACAACGAGAGATTTCCGACTATGCTAGATGCTGATATATACAGGTCTTAAT 720
721 AGCTTGAGAGGGAACAATGCGCAAGTTGGGTGCTTATAATCAATTCGCTAGAGATCTA 780
721 AGCTTGAGAGGGAACAATGCGCAAGTTGGGTGCTTATAATCAATTCGCTAGAGATCTA 780
781 ACCTTAGGGGTATAGATCTAGTGGCACTATCCCAAGCTATGCACTCGCACTTATCCA 840
781 ACCTTAGGGGTATAGATCTAGTGGCACTATCCCAAGCTATGCACTCGCACTTATCCA 840
841 ATAAATACAGTCTCAGTTTAAACAAGGGAAGTTTATACAGACCAATTTGGAGCAACAGG 900
841 ATAAATACAGTCTCAGTTTAAACAAGGGAAGTTTATACAGACCAATTTGGAGCAACAGG 900
901 GTAAATATGCAAGTATGAAATGCTATATATATATATATATATATATATATATATATAT 960
901 GTAAATATGCAAGTATGAAATGCTATATATATATATATATATATATATATATATATAT 960
961 ACTGCGGTATCCGAAGCCCGCATCTACTGATTTTCTAGAACAACTTAACTTATAGC 1020
961 ACTGCGGTATCCGAAGCCCGCATCTACTGATTTTCTAGAACAACTTAACTTATAGC 1020
1021 ACTTATACAGTGGTGTCTACTAGGCATATGACTTACTTGGGCGGGGCAACAATTTCAA 1080
1021 ACTTATACAGTGGTGTCTACTAGGCATATGACTTACTTGGGCGGGGCAACAATTTCAA 1080
1081 TCTCGGCAATAGAGGCGGATTAATATACCTCAACGATGCTTACCAATCTTCTATT 1140
1081 TCTCGGCAATAGAGGCGGATTAATATACCTCAACGATGCTTACCAATCTTCTATT 1140
1141 AATCCTGTAAAGATTAATCAATCTTCTCGAGAGCTATATATGGAATCAATATAGCAGGA 1200
1141 AATCCTGTAAAGATTAATCAATCTTCTCGAGAGCTATATATGGAATCAATATAGCAGGA 1200
1201 GTGCTTCTATGGGGAATTTACCTTGAACCTTATCATGCTGCTCTTCTTATAGATTAAAT 1260

1201 GTGCTTCTATGGGGAATTTACCTTGAACCTTATCATGCTGCTCCTACTGTTAGATTAAAT 1260
1261 TTTTAGGAACCTCAGAAATACCTTTTGAAGAGGTACTGCTAACTATAGTCAACCCCTATGAG 1320
1261 TTTTAGGAACCTCAGAAATACCTTTTGAAGAGGTACTGCTAACTATAGTCAACCCCTATGAG 1320
1321 TCACCTGGGCTTCAATTTAAAGATTGAGAACTGAAATTCACCAAGAAACCAACAGGCA 1380
1321 TCACCTGGGCTTCAATTTAAAGATTGAGAACTGAAATTCACCAAGAAACCAACAGGCA 1380
1381 CCAATTTATGAATCATATAGTCACTAGTATCTCAGATAGGGCTCATTTCAATCTAGG 1440
1381 CCAATTTATGAATCATATAGTCACTAGTATCTCAGATAGGGCTCATTTCAATCTAGG 1440
1441 GTGCATGTACCAAGTATATTTCTTGGACGACCGTAGTCAGATCGTACAAATACCAATTAGT 1500
1441 GTGCATGTACCAAGTATATTTCTTGGACGACCGTAGTCAGATCGTACAAATACCAATTAGT 1500
1501 TCGATAGCATAAACAAATACCAATTTGGTAAATCATTTCAACCTTAAATTCAGGTACCTCT 1560
1501 TCGATAGCATAAACAAATACCAATTTGGTAAATCATTTCAACCTTAAATTCAGGTACCTCT 1560
1561 GTAGTCAGTGGCCCGCAGGATTTTACAGGAGGGGATATATCCGAACTTAACGTTAATCGTAGT 1620
1561 GTAGTCAGTGGCCCGCAGGATTTTACAGGAGGGGATATATCCGAACTTAACGTTAATCGTAGT 1620
1621 GTACTAGTATGGGTCTTTAAATTTTAAATAATCATCATTAACGCGTATCGCGTAGAGATT 1680
1621 GTACTAGTATGGGTCTTTAAATTTTAAATAATCATCATTAACGCGTATCGCGTAGAGATT 1680
1681 CGTTATGCTGCTTCTCAAAACAATGGTCTGAGGGTAACTGTGCGAGGGAGTACTACTTTT 1740
1681 CGTTATGCTGCTTCTCAAAACAATGGTCTGAGGGTAACTGTGCGAGGGAGTACTACTTTT 1740
1741 GATCAAGGATTCCTTAGTACTAGTCAAAATGATGCTTTTGACATCTCAATCATTTAGA 1800
1741 GATCAAGGATTCCTTAGTACTAGTCAAAATGATGCTTTTGACATCTCAATCATTTAGA 1800
1801 TTTGCGAAGTTTCTGTTAGTATGTCATCTGCGAGTCAAACTGCTGGAATAAGTATA 1860
1801 TTTGCGAAGTTTCTGTTAGTATGTCATCTGCGAGTCAAACTGCTGGAATAAGTATA 1860
1861 AGTAAATATGCAAGTATGACAAACGTTTCTTTGATAAATTTGAATTTCAATTTCAATTTACT 1920
1861 AGTAAATATGCAAGTATGACAAACGTTTCTTTGATAAATTTGAATTTCAATTTCAATTTACT 1920
1921 GCAACCTTTCGAGCAGAAATACGATTTAGAAAGGCGCAAGAGCGGTGAATCTCTGTTT 1980
1921 GCAACCTTTCGAGCAGAAATACGATTTAGAAAGGCGCAAGAGCGGTGAATCTCTGTTT 1980
1981 ACTAATACGAAATTCGAAGAGATTGAAACAGATGTGACAGATTATCATATTCATCAAGTA 2040
1981 ACTAATACGAAATTCGAAGAGATTGAAACAGATGTGACAGATTATCATATTCATCAAGTA 2040
2041 TCCAAATTTAGTGGCGTGTATTCGGAATGAAATTTCTGCTTAGATGAAAGAGAGAAATTTACTT 2100
2041 TCCAAATTTAGTGGCGTGTATTCGGAATGAAATTTCTGCTTAGATGAAAGAGAGAAATTTACTT 2100
2101 GAGAAAGTGAATATGCGAAACGACTCAGTGTGATGAAAGAACTTACTCCAAAGATCCAAAC 2160
2101 GAGAAAGTGAATATGCGAAACGACTCAGTGTGATGAAAGAACTTACTCCAAAGATCCAAAC 2160
2161 TTTCAATCCATCAATAGGCAACAGACTTCTATCTACTTAAATGAGCAATCGAAATTTTACA 2220
2161 TTTCAATCCATCAATAGGCAACAGACTTCTATCTACTTAAATGAGCAATCGAAATTTTACA 2220
2221 TCTATCCATGAACAAATCTGAAATGATGGTGGGGAAGTGAGAACATTACAAATCCAGGAA 2280
2221 TCTATCCATGAACAAATCTGAAATGATGGTGGGGAAGTGAGAACATTACAAATCCAGGAA 2280
2281 GGAATGAGCTATTTAAAGAGAAATTAACGTCACATCCGCGGACTTTTAAATGAGTGTAT 2340
2281 GGAATGAGCTATTTAAAGAGAAATTAACGTCACATCCGCGGACTTTTAAATGAGTGTAT 2340

This Page Blank (uspto)

```
QY 2341 CCAGCGTATTTATATCAAAAATAGGAGCTCGGAATTAAGAAGCTTATCTCGCTACCA 2400
Db 2341 CCAGCGTATTTATCAAAAATAGGAGCTCGGAATTAAGAAGCTTATCTCGCTACCA 2400
QY 2401 TTAAGAGGGTATATTGAAGATAGTCAAGATTTAGAGATATATTGATTCGTTATAATGCG 2460
Db 2401 TTAAGAGGGTATATTGAAGATAGTCAAGATTTAGAGATATATTGATTCGTTATAATGCG 2460
QY 2461 AAACATGAACCATTTGGATGTTCCAGGTACCGAGTCCGATGCGGCTTTTCAGTTGAAGC 2520
Db 2461 AAACATGAACCATTTGGATGTTCCAGGTACCGAGTCCGATGCGGCTTTTCAGTTGAAGC 2520
QY 2521 CCAATCGGAGGTGCGGAGAACCGAATCGATGCGCACCATTTTGAATGGATCTGAT 2580
Db 2521 CCAATCGGAGGTGCGGAGAACCGAATCGATGCGCACCATTTTGAATGGATCTGAT 2580
QY 2581 CTAGATGTTCTCTGACAGATGAGAAAAATGCGCATCATTTCCCATCATTTCTCTTTG 2640
Db 2581 CTAGATGTTCTCTGACAGATGAGAAAAATGCGCATCATTTCCCATCATTTCTCTTTG 2640
QY 2641 GATATTGATATTGGATGTCACAGACTTGCATGAGAACTTAGGCGTGTGGTGTATTCAAG 2700
Db 2641 GATATTGATATTGGATGTCACAGACTTGCATGAGAACTTAGGCGTGTGGTGTATTCAAG 2700
QY 2701 ATTAAGACCGAGGATGTCATGCAAGACTAGGAATCTGGAATTTATTGAAGAGAAACCA 2760
Db 2701 ATTAAGACCGAGGATGTCATGCAAGACTAGGAATCTGGAATTTATTGAAGAGAAACCA 2760
QY 2761 TTATTAGGAGAACCATCTGCTCGTGTGAAGAGAGCAGAAAAATGGAGAGACAAACGT 2820
Db 2761 TTATTAGGAGAACCATCTGCTCGTGTGAAGAGAGCAGAAAAATGGAGAGACAAACGT 2820
QY 2821 GAAAAACTCAATTGGAACAAACAGGATATATACAGAGCAGAAAGCTGTGGATGCT 2880
Db 2821 GAAAAACTCAATTGGAACAAACAGGATATATACAGAGCAGAAAGCTGTGGATGCT 2880
QY 2881 TTATTGTAGATCTCAATATATAGATTACAAGCGGATACAAACATTTGGCATGATCAT 2940
Db 2881 TTATTGTAGATCTCAATATATAGATTACAAGCGGATACAAACATTTGGCATGATCAT 2940
QY 2941 GCGGCAGATAAATCTGTTCAATCGAATTCGAGAGGCTTATCTGTGAGAAATTTCTGTATC 3000
Db 2941 GCGGCAGATAAATCTGTTCAATCGAATTCGAGAGGCTTATCTGTGAGAAATTTCTGTATC 3000
QY 3001 CCGGGTGAATTCGGAATTTTGAAGAAATTAAGAGGTGCGATTTACTGCAATCTCC 3060
Db 3001 CCGGGTGAATTCGGAATTTTGAAGAAATTAAGAGGTGCGATTTACTGCAATCTCC 3060
QY 3061 CTATACGATCGGAGAAATGTCGTTAAAAATGGTGATTTTAAATAGGATTAAGCATGCTGG 3120
Db 3061 CTATACGATCGGAGAAATGTCGTTAAAAATGGTGATTTTAAATAGGATTAAGCATGCTGG 3120
QY 3121 AATGTAAGAGGCGATGATGATGTAACAGAGCCATCAGGTTCTGCTGTTATCCCA 3180
Db 3121 AATGTAAGAGGCGATGATGATGTAACAGAGCCATCAGGTTCTGCTGTTATCCCA 3180
QY 3181 GAATGGAAGCAGAACTGTAACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240
Db 3181 GAATGGAAGCAGAACTGTAACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240
QY 3241 CCGTGCACAGGCTACAAAGAGGATATGAGAGGGTGTGTAAACGATCCATGAATCGAG 3300
Db 3241 CCGTGCACAGGCTACAAAGAGGATATGAGAGGGTGTGTAAACGATCCATGAATCGAG 3300
QY 3301 AACATACAGAGCAGAACTTAAATTTGAAGAGGATGAGAGAGGATGATCAACGAT 3360
Db 3301 AACATACAGAGCAGAACTTAAATTTGAAGAGGATGAGAGAGGATGATCAACGAT 3360
QY 3361 ACAGGAACTGTGAATGATTATCTGCACACCAAGGTACAGAGGATGATTAATCCCGTAT 3420
Db 3361 ACAGGAACTGTGAATGATTATCTGCACACCAAGGTACAGAGGATGATTAATCCCGTAT 3420
```

```
QY 3421 GCTGGATATGAGATGCATATGAAGTTGATCTACAGCATCTGTTAATTACAACCGACT 3480
Db 3421 GCTGGATATGAGATGCATATGAAGTTGATCTACAGCATCTGTTAATTACAACCGACT 3480
QY 3481 TATGAAGAAAGAAACGTTATACAGATGTACGAAGAGATAATCATTTGTGAATATGACAGAGG 3540
Db 3481 TATGAAGAAAGAAACGTTATACAGATGTACGAAGAGATAATCATTTGTGAATATGACAGAGG 3540
QY 3541 TATGTGAATTTCCACCATCTACAGCTGTTATATATGACAAAAAGAAATTAGAATATCTCCCA 3600
Db 3541 TATGTGAATTTCCACCATCTACAGCTGTTATATATGACAAAAAGAAATTAGAATATTTCCCA 3600
QY 3601 GAAACCGATAAGGTATGATTCGATTCGATTCGAGAAACGGAAGGAACTTCTATGTCAGCAGC 3660
Db 3601 GAAACCGATAAGGTATGATTCGATTCGATTCGAGAAACGGAAGGAACTTCTATGTCAGCAGC 3660
QY 3661 GTGGAACTACTCTTATGAGGAATAG 3687
Db 3661 ATAGAACTACTCTTATGAGGAATAG 3687

RESULT 3
US-10-428-961-62
; Sequence 62, Application US/10428961
; Publication No. US20030237111A1
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Rupar, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin
; FILE REFERENCE: Polynucleotides, Compositions, and Methods of Use (Amended)
; CURRENT FILING DATE: 2003-05-02
; PRIOR FILING DATE: 2003-05-02
; PRIOR FILING DATE: 2000-09-13
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 62
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-10-428-961-62

Query Match 88.4%; Score 3260.2; DB 17; Length 3684;
Best Local Similarity 93.3%; Pred. No. 0;
Matches 3445; Conservative 0; Mismatches 233; Indels 15; Gaps 3;

QY 1 TTGACTTCAATAGAAAAATGGAATGAAATATAAATGCTTTATCGATTCGAGCTGA 60
Db 1 TTGACTTCAATAGAAAAATGGAATGAAATATAAATGCTTTATCGATTCGAGCTGA 60
QY 61 TCGAATCATTCACACAAATGATCTATCACCAGATGCTCGTATTGAGGATCTTTGTGT 120
Db 61 TCGAATCATTCGACAAATGATCTATCACCAGATGCTCGTATTGAGGATGATGTTGT 120
QY 121 ATAGCCGAGGGGAATATATCAATCCATTGTTAGCGGATCAACAGTCCAAACGGGTATT 180
Db 121 ATAGCCGAGGGGAATATATCAATCCATTGTTAGCGGATCAACAGTCCAAACGGGTATT 180
QY 181 AACATAGCTGTAGATAGTACTAGGTGATTTAGCGGTACCGTTTCTGCGGACAAATAGTAGT 240
Db 181 AACATAGCTGTAGATAGTACTAGGTGATTTAGCGGTACCGTTTCTGCGGACAAATAGTAGT 240
QY 241 TTTTATAGTTTCTTGTGTTGGAATTTAGCGGCGCGGACAGAGATCAGTGGGAAATTTTC 300
Db 241 TTTTATAGTTTCTTGTGTTGGAATTTAGCGGCGCGGACAGAGATCCTTGGGAAATTTTC 300
QY 301 CTAGAACATGTCGAACAACTTTATAAATCAACAAATAACAGAAAAATGTAGGAATACGGCA 360
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2005, 03:37:54 ; Search time 58 Seconds
(without alignments)
1580.501 Million cell updates/sec

Title: US-10-614-524-2
Perfect score: 6479
Sequence: 1 LTSNRKNEIINALSIPAV.....IGTEGKFIVDSVELLMEE 1228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6462	99.7	1228	4	US-09-661-322A-38
2	5926.5	91.5	1227	1	US-08-448-170-8
3	5926.5	91.5	1227	3	US-08-961-803-9
4	5912.5	91.3	1227	4	US-09-661-322A-63
5	5659.5	87.4	1207	1	US-07-951-715A-7
6	5659.5	87.4	1207	2	US-08-459-448A-7
7	5659.5	87.4	1207	3	US-08-459-595A-7
8	5659.5	87.4	1207	3	US-08-459-504B-7
9	5659.5	87.4	1207	3	US-08-459-444-7
10	5659.5	87.4	1207	3	US-09-053-549-8
11	5659.5	87.4	1207	3	US-09-547-422-7
12	5659.5	87.4	1207	3	US-09-988-462-7
13	5436.5	83.9	1227	3	US-09-053-549-2
14	5237.5	80.8	1229	1	US-08-100-709-4
15	5237.5	80.8	1229	1	US-08-176-865-4
16	5237.5	80.8	1229	1	US-08-474-038-4
17	5237.5	80.8	1229	2	US-08-779-046-4
18	5237.5	80.8	1229	2	US-08-881-340-4
19	5108	78.8	1186	3	US-09-178-252-23
20	5108	78.8	1186	4	US-09-826-660-23
21	3977.5	61.4	1174	1	US-08-040-751-3
22	3977.5	61.4	1174	1	US-08-291-368-2
23	3977.5	61.4	1174	2	US-08-962-190-2
24	3977.5	61.4	1174	5	PCT-US95-10310-2
25	3977.5	61.4	1174	6	5164180-4
26	3977.5	61.4	1174	6	5164180-4
27	3791	58.5	1176	1	US-08-434-823-2

28	3791	58.5	1176	1	US-08-457-366-2	Sequence 2, Appli
29	3579	55.2	1170	1	US-08-032-364-2	Sequence 2, Appli
30	3556	54.9	1167	1	US-08-100-709-2	Sequence 2, Appli
31	3556	54.9	1167	1	US-08-176-865-2	Sequence 2, Appli
32	3556	54.9	1167	1	US-08-474-038-2	Sequence 2, Appli
33	3556	54.9	1167	2	US-08-779-046-2	Sequence 2, Appli
34	3556	54.9	1167	2	US-08-881-340-2	Sequence 2, Appli
35	3502.5	54.1	1189	2	US-08-980-071-59	Sequence 59, Appl
36	3502.5	54.1	1189	3	US-09-314-093-59	Sequence 59, Appl
37	3502.5	54.1	1189	3	US-09-337-635-59	Sequence 59, Appl
38	3502.5	54.1	1189	4	US-09-337-280-59	Sequence 59, Appl
39	3502.5	54.1	1189	4	US-09-972-175-59	Sequence 59, Appl
40	3502.5	54.1	1189	4	US-10-200-522-59	Sequence 59, Appl
41	3500.5	54.0	1189	2	US-08-980-071-2	Sequence 2, Appli
42	3500.5	54.0	1189	2	US-08-757-536-2	Sequence 2, Appli
43	3500.5	54.0	1189	3	US-09-314-093-2	Sequence 2, Appli
44	3500.5	54.0	1189	3	US-09-250-848-2	Sequence 2, Appli
45	3500.5	54.0	1189	3	US-09-251-885-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-661-322A-38
; Sequence 38, Application US/09661322A
; Patent No. 6593293

; GENERAL INFORMATION:

; APPLICANT: Baum, James A.

; APPLICANT: Chu, Chih-Rei

; APPLICANT: Donovan, William P.

; APPLICANT: Gilmer, Amy J.

; APPLICANT: Rupar, Mark J.

; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin Composi

; TITLE OF INVENTION: and Methods of Use

; FILE REFERENCE: MECO201

; CURRENT APPLICATION NUMBER: US/09/661,322A

; CURRENT FILING DATE: 2000-09-13

; NUMBER OF SEQ ID NOS: 63

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 38

; LENGTH: 1228

; TYPE: PRT

; ORGANISM: Bacillus thuringiensis

US-09-661-322A-38

Query Match 99.7%; Score 6462; DB 4; Length 1228;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1223; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY	1	LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI	60
Db	1	MTSNRKNNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI	60
QY	61	NIAGRILGVLPFAGQIASFYSLGELWPRGRDQWEIFLEHVEQLINQOITENARNTA	120
Db	61	NIAGRILGVLPFAGQIASFYSLGELWPRGRDQWEIFLEHVEQLINQOITENARNTA	120
QY	121	LARLOGGDSFRAYQOSLEDWLENRDDARTSRVLYTQYIALELDFLNAPLFAIRNOEVP	180
Db	121	LARLOGGDSFRAYQOSLEDWLENRDDARTSRVLYTQYIALELDFLNAPLFAIRNOEVP	180
QY	181	LLMVAQAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVEQTRDYSYCVWEYNTGLN	240
Db	181	LLMVAQAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVEQTRDYSYCVWEYNTGLN	240
QY	241	SLRGTTNAASWVRYNQFRDRTLGLVLDLVALFPSTYDTRTYPINTSAQLTRVYTDATGATG	300
Db	241	SLRGTTNAASWVRYNQFRDRTLGLVLDLVALFPSTYDTRTYPINTSAQLTRVYTDATGATG	300
QY	301	VNMASMMWNNAPSPSAITAVIRSPHLLDFLEQLTIFSTSSRWSATRMTYWRGHTIQ	360
Db	301	VNMASMMWNNAPSPSAITAVIRSPHLLDFLEQLTIFSTSSRWSATRMTYWRGHTIQ	360

```

361 SRPIGGGLNTSTHGSTNTSINPVLRSFSDVYVWTSYAGVLLMGVLEPIHGVPTVRFN 420
361 SRPIGGGLNTSTHGSTNTSINPVLRSFSDVYVWTSYAGVLLMGVLEPIHGVPTVRFN 420
421 FRNPQTFERTGANYSPYSPGQLKDSSTETLPPETTERPNYESYSHRLSHIGLISQSR 480
421 FRNPQTFERTGANYSPYSPGQLKDSSTETLPPETTERPNYESYSHRLSHIGLISQSR 480
481 VHVYVSWTHRSADRTNTISSDSTQIPLVKSFNLSGTSVSGPGFTGGDIIRTNVGS 540
481 VHVYVSWTHRSADRTNTISSDSTQIPLVKSFNLSGTSVSGPGFTGGDIIRTNVGS 540
541 VLSMGLNFNNTSLQRYVRVRYAASQTMVLRTVVGSTTFDQGPSTMSANESLTSQSPR 600
541 VLSMGLNFNNTSLQRYVRVRYAASQTMVLRTVVGSTTFDQGPSTMSANESLTSQSPR 600
601 FAEFPVGISASGQTAGISINNAGRTQTFHDKIEPIPTATFEAEYDLERAQEAVALNF 660
601 FAEFPVGISASGQTAGISINNAGRTQTFHDKIEPIPTATFEAEYDLERAQEAVALNF 660
661 TTNTPRLKTDVTDYHIDQVSNLVACLSDEFCLDEKRELLEKVKYAKRLSDERNLLQDPN 720
661 TTNTPRLKTDVTDYHIDQVSNLVACLSDEFCLDEKRELLEKVKYAKRLSDERNLLQDPN 720
721 FTSINKOPDFISTNEQSNFTSIHQSEHGWGSENITIQEGNDVFKENYVTLPGTFNECY 780
721 FTSINKOPDFISTNEQSNFTSIHQSEHGWGSENITIQEGNDVFKENYVTLPGTFNECY 780
781 PTLYQKIGESLKAQTRYQIRGYEDSQLEIYLIRYNAKHETLDVPGTESVWPLSVES 840
781 PTLYQKIGESLKAQTRYQIRGYEDSQLEIYLIRYNAKHETLDVPGTESVWPLSVES 840
841 PIGRCGEPRNCAHPFNWPLDSCRDGCKCAHSHHFSLDIDIGCTDLHENLGVWVVF 900
841 PIGRCGEPRNCAHPFNWPLDSCRDGCKCAHSHHFSLDIDIGCTDLHENLGVWVVF 900
901 IKTOEGHARLGNLEFTEKPLLGALSRVRAKRWDRKQLETKRVYTTAKAVALDA 960
901 IKTOEGHARLGNLEFTEKPLLGALSRVRAKRWDRKQLETKRVYTTAKAVALDA 960
961 LFVDSQVNRLOADNTIGMHAADKLVRHIREAYLSLSVIPGVNAEIFELEGRITTAIS 1020
961 LFVDSQVNRLOADNTIGMHAADKLVRHIREAYLSLSVIPGVNAEIFELEGRITTAIS 1020
1021 LYDARNVVKNGDFNGLACNVKGVHDVQOQSHRSVLVPEWEAEVSQAVRVCGRGYIL 1080
1021 LYDARNVVKNGDFNGLACNVKGVHDVQOQSHRSVLVPEWEAEVSQAVRVCGRGYIL 1080
1081 RVTAYKEGYGEGCVTTHEIENNTDELKFKNCEEEVYPTDTGTCNDYTAHOGTAVCNRSN 1140
1081 RVTAYKEGYGEGCVTTHEIENNTDELKFKNCEEEVYPTDTGTCNDYTAHOGTAVCNRSN 1140
1141 AGYEDAYEDVTASVNYKPYEETTYTDVRRDNHCEYDRGVNYPPLPAGYMTKELEYFP 1200
1141 AGYEDAYEDVTASVNYKPYEETTYTDVRRDNHCEYDRGVNYPPLPAGYMTKELEYFP 1200
1201 ETDKWIEIGETGKFIIVDSVELLMBE 1228
1201 ETDKWIEIGETGKFIIVDSVELLMBE 1228

```

RESULT 2

```

US-08-448-170-8
; Sequence 8, Application US/08448170
; Patent No. 5723758
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel
; APPLICANT: Cummings, David A.
; APPLICANT: Cannon, Raymond J.C.
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stelman, Steve
; TITLE OF INVENTION: No. 5723758el Bacillus thuringiensis Isolate Denoted

```

```

; TITLE OF INVENTION: B.t. PS158C2, Active Against Lepidopteran Pests, and Genes
; TITLE OF INVENTION: Encoding Lepidopteran-Active Toxins
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,170
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/069,902
; FILING DATE: 01-JUNE-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/759,247
; FILING DATE: 13-SEPT-1991
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/S 102D.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904) 375-8100
; TELEFAX: (904) 372-5800
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1227 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-448-170-8

```

```

Query Match 91.5%; Score 5926.5; DB 1; Length 1227;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 1129; Conservative 36; Mismatches 60; Indels 5; Gaps 3;

QY 1 LTSNRKNEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db 1 LTSNRKNEIINALSIPAVSNHSAQNLSTDAIEDSLCIAEGNNIDPPVSASTVQTGI 60

QY 61 NIAGRILGLVGPFAQGIASFYFLVGLWPRGRDQWEIFLEHVEQLINQOITENARNTA 120
Db 61 NIAGRILGLVGPFAQGIASFYFLVGLWPRGRDQWEIFLEHVEQLIRQOVTENTRDTA 120

QY 121 LARLOGLDPSFRAYQOSLEDWLENRDDARTSRVLYTQYIALELDFLNAPLFAIRNOEVP 180
Db 121 LARLOGLDPSFRAYQOSLEDWLENRDDARTSRVLYTQYIALELDFLNAPLFAIRNOEVP 180

QY 181 LLMVYAAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVEQTRDYSYCVWEYNTGLN 240
Db 181 LLMVYAAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVEQTRDYSYCVWEYNTGLN 240

QY 241 SLRGTAASWVRYNQRRDLTLGLVLDLVALFSDYDTRTYPTINTSAQITREYVYDAIGATG 300
Db 241 NLRGTAESWLRYNQRRDLTLGLVLDLVALFSDYDTRTYPTINTSAQITREYVYDAIGATG 300

QY 301 V--NMASMMWYNNAPSAIETAVIRSPHLLDLFQLTIFSTSSWSATRHMTYWRGHT 358
Db 301 APSGPASTWFWNNAPSAIENAVIRPPHLLDFPQLTIFSVLRWSNTQYNNYVWGR 360

QY 359 IQSRPIGGGLNTSTHGSTNTSINPVLRSFSDVYVWTSYAGVLLMGVLEPIHGVPTVRFN 418
Db 359 IQSRPIGGGLNTSTHGSTNTSINPVLRSFSDVYVWTSYAGVLLMGVLEPIHGVPTVRFN 418

```

```
Db 361 LESRTIRGSLSTSTHGTNTSINPVTQLFTSRDVRTSTPAGINI--LLTTTPVNGVPEWAR 418
Qy 419 FNFNPQNTFPGTANYSQPYESQGLQLOKSETLPPETTERPNYESYSHRLSHIGLISQ 478
Db 419 FNFNPLNSL-RGSLLYTIGYTGVTQLFDSSETLPPETTERPNYESYSHRLSIRLISG 477
Qy 479 SRVHPVYVSWTHRADRTNTSSDSTIQIPLVKFNLNSGTSVVGPGFTGGDIIRTNVN 538
Db 478 NTLRAPYVSWTHRADRTNTSSDSTIQIPLVKFNLNSGTSVVGPGFTGGDIIRTNVN 537
Qy 539 GSVLSMGLNFNTSLQRYRVRVRYAASQTMVLRVTGSGSTTFDQGFSTMSANESLSQS 598
Db 538 GSVLSMGLNFNTSLQRYRVRVRYAASQTMVLRVTGSGSTTFDQGFSTMSANESLSQS 597
Qy 599 FRFAEPFVGISASGSGTAGISISNNAGROTQPHFKIEFIPITATFEAYDYLERAOEAVNA 658
Db 598 FRFAEPFVGISASGSGTAGISISNNAGROTQPHFKIEFIPITATFEAYDYLERAOEAVNA 657
Qy 659 LFTNTNPRRLKTDVTDHIDOVSNLVACLDFECLDEKRELLKVKYAKRLSDERNLLQD 718
Db 658 LFTNTNPRRLKTDVTDHIDOVSNLVACLDFECLDEKRELLKVKYAKRLSDERNLLQD 717
Qy 719 PNFTSINKQPDFISTNQSNFTSHEQSEHGWMGSENITIQEGNDVFKENVYVILPGTFNE 778
Db 718 PNFTSINKQPDFISTNQSNFTSHEQSEHGWMGSENITIQEGNDVFKENVYVILPGTFNE 777
Qy 779 CYPTLYQKIGESLKYATRYQLRGYIEDSODLEIYLIRYNNAKHETLDVPGTESVWPLSV 838
Db 778 CYPTLYQKIGESLKYATRYQLRGYIEDSODLEIYLIRYNNAKHETLDVPGTESVWPLSV 837
Qy 839 ESPIGRCGEPNRCAPHFEPWPDLDSCRDGKCAHSHHPSLDDIDGCTDLHENLGVWV 898
Db 838 ESPIGRCGEPNRCAPHFEPWPDLDSCRDGKCAHSHHPSLDDIDGCTDLHENLGVWV 897
Qy 899 FKIKTOGHARLGNLEFTEEPKPLGEALSRVKRAKKWRDKREKLQLETKRVYTEAKEAV 958
Db 898 FKIKTOGHARLGNLEFTEEPKPLGEALSRVKRAKKWRDKREKLQLETKRVYTEAKEAV 957
Qy 959 DALFVDSQYNRLOADNTNIGMHAADKLVHRIREAYLSLSVPGVNAEIPFEELEGRITTA 1018
Db 958 DALFVDSQYNRLOADNTNIGMHAADKLVHRIREAYLSLSVPGVNAEIPFEELEGRITTA 1017
Qy 1019 ISLYDARNVKNKGFNNGLACWNVKGVHVDVQOSSHRSVLVPIPEWAEVSQAVRVCPRGY 1078
Db 1018 ISLYDARNVKNKGFNNGLACWNVKGVHVDVQOSSHRSVLVPIPEWAEVSQAVRVCPRGY 1077
Qy 1079 ILRVTAYKEGEGECVTHETENNIDELKFNKCEEEVYPTDCTCNDYTAHQGTAVCNS 1138
Db 1078 ILRVTAYKEGEGECVTHETENNIDELKFNKCEEEVYPTDCTCNDYTAHQGTAAACS 1137
Qy 1139 RNAGYEDAYEVDTTASVNYKPYTEETVTVDRRDHCEYDRGYVNYPPPLPAGYMTKELEY 1198
Db 1138 RNAGYEDAYEVDTTASVNYKPYTEETVTVDRRDHCEYDRGYVNYPPVPGYMTKELEY 1197
Qy 1199 FPETDKWIEIGETEGKFIIVDSVLELLMEE 1228
Db 1198 FPETDKWIEIGETEGKFIIVDSVLELLMEE 1227
```

RESULT 3

```
US-08-961-803-9
; Sequence 9, Application US/08961803
; Patent No. 6150589
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel
; APPLICANT: Cummings, David A.
; APPLICANT: Cannon, Raymond J.C.
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stelman, Steve
; TITLE OF INVENTION: No. 6150589a1 Bacillus thuringiensis Isolate Denoted
; TITLE OF INVENTION: B.t. PS158C3, Active Against Lepidopteran Pests, and Genes
; TITLE OF INVENTION: Encoding Lepidopteran-Active Toxins
; NUMBER OF SEQUENCES: 10
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jay M. Sanders
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,803
; FILING DATE: 31-OCT-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/069,902
; FILING DATE: 01-JUNE-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/759,247
; FILING DATE: 13-SEPT-1991
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/448,170
; FILING DATE: 23-MAY-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: M/S 102DCD1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (352) 375-8100
; TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1227 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-961-803-9
```

Query Match 91.5%; Score 5926.5; DB 3; Length 1227;

Best Local Similarity 91.8%; Pred. No. 0;

Matches 1129; Conservative 36; Mismatches 60; Indels 5; Gaps 3;

```
Qy 1 LTSNRKNEEIIINALSTIPAVSNHSTQMDLSPDARIEDSLCIAEGNINPLVSASTVQTGI 60
Db 1 LTSNRKNEEIIINALSTIPAVSNHSAQMNLSLTDARIEDSLCIAEGNIDPPFVSASTVQTGI 60
Qy 61 NIAGRILGLGVFPAGQIASFYSLVGLWPRGRDOMEIFLEHVQQLINOQITENARNTA 120
Db 61 NIAGRILGLGVFPAGQIASFYSLVGLWPRGRDPWEIFLEHVQQLIRQOVTENTRDTA 120
Qy 121 LARLQGLGDSFRAYQQSLEDWLENRDDARTSRVLYTQYIALELDLFLNAPLFAIRNOEVP 180
Db 121 LARLQGLGNSFRAYQQSLEDWLENRDDARTSRVLYTQYIALELDLFLNAPLFAIRNOEVP 180
Qy 181 LLMVYQAANLHLLLDASLFGSEFGLTSQETQRYRQVEQTRDYSDYCVHYNTGLN 240
Db 181 LLMVYQAANLHLLLDASLFGSEFGLTSQETQRYRQVEQTRDYSDYCARWYNTGLN 240
Qy 241 SLRGTHAASVVRVYNOPRDLTLGVLDLVALPPSDVDTETYPINTSAQLTREVYTDAGTG 300
Db 241 NLRGTHAASVVRVYNOPRDLTLGVLDLVALPPSDVDTETYPINTSAQLTREVYTDAGTG 300
Qy 301 V--NMASMNWYNNNAPSFAIETAVIRSPHLLDFLEQLTIFSTSSRWSATRHMTYWRGHT 358
Db 301 APSGFASTNFWNNAPSFSAIEAAVIRPPHLLDFPEQLTIFSVLSRWSNTQYMWYWGHR 360
Qy 359 IQSRPIGGGLNTSTHGSTNTSINPVRLSFFPSRDVYWTESVAGVLLWGIYLEPIHGVTVR 418
```



```

Qy 899 PKIKTQCHARKLGNLEFIEEKPLLGALSRVKRAEKKWRDKREKQLQLETKRVYTEAKEAV 958
Db 898 PKIKTQCHARKLGNLEFIEEKPLLGALSRVKRAEKKWRDKREKQLQLETKRVYTEAKEAV 957
Qy 959 DALFVDSQYNRLQADTNGIMHAADKLVRIRREAYLSLSVPCVNAEIEEELGRIITA 1018
Db 958 DALFVDSQYDKLQADTNGIMHAADKLVRIRREAYLSLSVPCVNAEIEEELGRIITA 1017
Qy 1019 ISLYDARNVKNQDFNGLACWNVKGVHVDVQOSSHRSVLVIPWEAEVSAQVRVCPGRGY 1078
Db 1018 ISLYDARNVKNQDFNGLACWNVKGVHVDVQOSSHRSVLVIPWEAEVSAQVRVCPGRGY 1077
Qy 1079 ILRVATYKEGEGECVTIHEIENNTDELKFKNCBEEVYPTDTCNDYTAHQGTAVCNS 1138
Db 1078 ILRVATYKEGEGECVTIHEIENNTDELKFKNCBEEVYPTDTCNDYTAHQGTAVCNS 1137
Qy 1139 RNAGYEDAYEVDTTASVNYKPYEEETVTVRRDNHCEYDRGVYNYPPVAGYMTKSELEY 1198
Db 1138 RNAGYEDAYEVDTTASVNYKPYEEETVTVRRDNHCEYDRGVYNYPPVAGYMTKSELEY 1197
Qy 1199 FPETDKWIEIGETEGKFIIVDSVELLMEE 1228
Db 1198 FPETDKWIEIGETEGKFIIVDSVELLMEE 1227

```

RESULT 5

```

US-07-951-715A-7
; Sequence 7, Application US/07951715A
; Patent No. 5625136
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lytle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/951,715A
; FILING DATE: 25-SEP-1992
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: S-18805/A/CSC 1577/CIP

```

TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: (919)541-8615
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-951-715A-7

```

```

Query Match 87.4%; Score 5659.5; DB 1; Length 1207;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;

```

```

Qy 27 MDLSPDARIEDSLCIAEGNNINPLVSASTVQTGINIAGRILGVLPFAGQIASFYSLV 86
Db 1 MDLLPDARIEDSLCIAEGNNIDPPVSASTVQTGINIAGRILGVLPFAGQIASFYSLV 60
Qy 87 GELWPRGRDQWEIFLEHVEQLINQOITENARNALTALRQGLGDSFRAYQOSLEWLENRD 146
Db 61 GELWPRGRDQWEIFLEHVEQLINQOITENARNALTALRQGLGDSFRAYQOSLEWLENRD 120
Qy 147 DARTSRVLYTOYIALELDFLNAMPFLFAIRNOEVPFLMVYQAANLHLLLEDASLFSEF 206
Db 121 DARTSRVLYTOYIALELDFLNAMPFLFAIRNOEVPFLMVYQAANLHLLLEDASLFSEF 180
Qy 207 GLTSQEIQRVYEROVETROYSYCVWEYNTGLNSLRGTNAASWVRVYNOQFRDRLTLGVLD 266
Db 181 GLTSQEIQRVYEROVETROYSYCVWEYNTGLNSLRGTNAASWVRVYNOQFRDRLTLGVLD 240
Qy 267 LVALFPSYDTRTYPIINTSAQLTREYVYTDATGATGVNMAWNNYNNAPSAITAVIRS 326
Db 241 LVALFPSYDTRTYPIINTSAQLTREYVYTDATGATGVNMAWNNYNNAPSAITAVIRS 300
Qy 327 PHLLDFLEQLTIFSTSRWSATRHMTYWRGHTTQSRPIGGGLNTSTHGNTSINPVRLS 386
Db 301 PHLLDFLEQLTIFSTSRWSATRHMTYWRGHTTQSRPIGGGLNTSTHGNTSINPVTLR 360
Qy 387 FFSRDVYVWTSYAGVLLGWLYLEPIHGVPTVRNFRNPQNTFERGTANYSPYSPGLQL 446
Db 361 FASRDVYVWTSYAGVLLGWLYLEPIHGVPTVRNFRNPQNTFERGTANYSPYSPGLQL 420
Qy 447 KDSETELPETTERPNYESYSHRLSHIGLSQSRHVHPVYVSWTHRSADRTNTISDSITQ 506
Db 421 KDSETELPETTERPNYESYSHRLSHIGLSQSRHVHPVYVSWTHRSADRTNTIGPNRITQ 480
Qy 507 IPLVKSFNLSGTSVSGPGFTGDDIIRTNVNGSVLSMGLNFNTSLQRYRVRYAASQ 566
Db 481 IPVYKASELPQGTVVRGPGFTGDDILRLRTNTGCGPIRVTVNGPLTQRYRIGFRYASTV 540
Qy 567 TMVLRTVVGSTTFDQGFPTMSANESLTSQSRFAEPFVPGISASGSQ-TAGISISNAG 625
Db 541 DFDFFVSRGGTTVNNFRFLTMSGDDELKYNFVRRAFTPTPTQIQDIIRTSIQGLSG 600
Qy 626 RQTFHFDKIEFIPITATFEAEYDLERAQEAVALFTNTNPRRLKTDVTDYHIDQVSNLVA 685
Db 601 NGEVYIDKIEIIPVTATFEAEYDLERAQEAVALFTNTNPRRLKTDVTDYHIDQVSNLVA 660
Qy 686 CLSDEFCLDEKRELLEKVKYAKRLSDERNLLQDPNFTSINKQDPFITNQSOSNFTSHEQ 745
Db 661 CLSDEFCLDEKRELLEKVKYAKRLSDERNLLQDPNFTSINKQDPFITNQSOSNFTSHEQ 720
Qy 746 SEHGWSGSENIITQEGNDVFEKNYVTLPGTFNECYPTLYOKYGESELKAYTRQYLRGYI 805
Db 721 SEHGWSGSENIITQEGNDVFEKNYVTLPGTFNECYPTLYOKYGESELKAYTRQYLRGYI 780
Qy 806 EDSQDLIELYLRNAXKHETLDVPGTESVWPLSVESPIGRCEPNRCAPHFENWDLDCSC 865
Db 781 EDSQDLIELYLRNAXKHETLDVPGTESVWPLSVESPIGRCEPNRCAPHFENWDLDCSC 840
Qy 866 RDGEKCAHSHHFLSDIDICTDLHENLGVVWVFKINTQEGHARLGNLEFIEEKPLGEEA 925

```

841 RDGEKCAHSHHSFSLDIDVGCTDLHENLGVWVFKIQTBQCHARLGNLEFIEBKPLLGEA 900
 926 LSRVKRAEKWRDKREKLOLETKVYVTEAEAVDALFVDSQYNRLQADTWIGMHAADKL 985
 901 LSRVKRAEKWRDKREKLOLETKVYVTEAEAVDALFVDSQYNRLQADTWIGMHAADKL 960
 986 VHRIRAYLSLSVPGVNAEIPFEELGRITITAIISLYDARNVVKNGDFNNGLACVNVKGH 1045
 961 VHRIRAYLSLVPVPGVNAEIPFEELGHITITAIISLYDARNVVKNGDFNNGLTCVNVKGH 1020
 1046 VDVQOQSHRSVLVPEWEAEVSQAVRCPGRGYILRVYAYKEGEGCVTIHIEINNTDE 1105
 1021 VDVQOQSHRSVLVPEWEAEVSQAVRCPGRGYILRVYAYKEGEGCVTIHIEINNTDE 1080
 1106 LKFKNCEEEVYPTDGTCTNDYTAHOGTA---VCNRSNAGYEDAYEVDVTTASVNYKPTY 1161
 1081 LKFKNCEEEVYPTDGTCTNDYTAHOGTACACACNSNAGYEDAYEVDVTTASVNYKPTY 1140
 1162 EETTYTDVRDNDHCEYDRGVNYPPLPAGYMTKELEYFPETDKVWIEIGETEGKFIVDSV 1221
 1141 EETTYTDVRDNDHCEYDRGVNYPPLPAGYMTKELEYFPETDKVWIEIGETEGKFIVDSV 1200
 1222 ELLMEE 1228
 1201 ELLMEE 1207

RESULT 6

US-08-459-448A-7
 ; Sequence 7, Application US/08459448A
 ; Patent No. 5859336
 ; GENERAL INFORMATION:
 ; APPLICANT: Koziel, Michael G.
 ; APPLICANT: Desai, Nalini M.
 ; APPLICANT: Lewis, Kelly S.
 ; APPLICANT: Kramer, Vance C.
 ; APPLICANT: Warren, Gregory W.
 ; APPLICANT: Evola, Stephen V.
 ; APPLICANT: Crossland, Lyle D.
 ; APPLICANT: Wright, Martha S.
 ; APPLICANT: Launis, Karen L.
 ; APPLICANT: Rothstein, Steven J.
 ; APPLICANT: Bowman, Cindy G.
 ; APPLICANT: Dawson, John L.
 ; APPLICANT: Dunder, Erik M.
 ; APPLICANT: Pace, Gary M.
 ; APPLICANT: Suttie, Janet L.
 ; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
 ; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
 ; NUMBER OF SEQUENCES: 94
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 5859336artis Corporation
 ; STREET: Patent & Trademark Dept., 520 White Plains
 ; STREET: Rd., POB 2005
 ; CITY: Tarrytown
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10591-9005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patencin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/459,448A
 ; FILING DATE: 02-JUN-1995
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/951,715
 ; FILING DATE: 25-SEP-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/772,027

; FILING DATE: 04-OCT-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pace, Gary M.
 ; REGISTRATION NUMBER: 40403
 ; REFERENCE/DOCKET NUMBER: CGC 1577/CLIP/DIV4
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (919) 541-8582
 ; TELEFAX: (919) 541-8689
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1207 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-459-448A-7
 ;
 ; Query Match 87.4%; Score 5659.5; DB 2; Length 1207;
 ; Best Local Similarity 89.4%; Pred. No. 0;
 ; Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;
 ;
 QY 27 MDLSPDARIEDSLCIAEGNNINPLVSASTVOTGINIAGRIILGVLPFAGQIASFYSEFLV 86
 DB 1 MDLLPDARIEDSLCIAEGNNIDPFVSASTVOTGINIAGRIILGVLPFAGQIASFYSEFLV 60
 QY 87 GELWPRGRDQWEIFLEHVEQLINQOITENARNTALRQGLGDSFRAYQOSLEDWLENRD 146
 DB 61 GELWPRGRDQWEIFLEHVEQLINQOITENARNTALRQGLGDSFRAYQOSLEDWLENRD 120
 QY 147 DARTRSVLTYQYIALELDFINAMPLFAIRNQEVPLLMVYAQAANLHLLLRDASLFGSEF 206
 DB 121 DARTRSVLTYQYIALELDFINAMPLFAIRNQEVPLLMVYAQAANLHLLLRDASLFGSEF 180
 QY 207 GLTSQEIORYYERQVETRDYSDYCVWYNTGLNSRGNTAASWVRYNQPRRDLTLGVLD 266
 DB 181 GLTSQEIORYYERQVETRDYSDYCVWYNTGLNSRGNTAASWVRYNQPRRDLTLGVLD 240
 QY 267 LVALFPSYDTRTYPINTSAQLTREVTDAIGATGVNNAAMNWNNAAPSALETAVIRS 326
 DB 241 LVALFPSYDTRTYPINTSAQLTREVTDAIGATGVNNAAMNWNNAAPSALETAVIRS 300
 QY 327 PHLLDFLEQLTTFSTSRWSATRHMTYWRGHTIQSRPIGGGLNTSTHGSTNTSINPVRLS 386
 DB 301 PHLLDFLEQLTTFSTSRWSATRHMTYWRGHTIQSRPIGGGLNTSTHGSTNTSINPVRLS 360
 QY 387 FESRDVWVTSYAGVLLWGLYLEPIHGVTVRNPNPONTFERGTANTSQPVESPCQLQ 446
 DB 361 FASRDVWVTSYAGVLLWGLYLEPIHGVTVRNPNPONTFERGTANTSQPVESPCQLQ 420
 QY 447 KDSETELPPTETTERPNYESYSHRLSHIGLISOSRVHPVYVSWTHRSADRTNTSSDITQ 506
 DB 421 KDSETELPPTETTERPNYESYSHRLSHIGLISOSRVHPVYVSWTHRSADRTNTSSDITQ 480
 QY 507 IPLVKSFNLSGTSVVGPGFTGGDIIRTNVNGSVLSMGLNFNNTSLQRYRVVRVAASQ 566
 DB 481 IPMVKASELPQGTTVVRGPGFTGGDILRRNTTGGFPIRVTVNGPLTQRYRIGRYASTV 540
 QY 567 TMVLRVTVGSTTFDQGFSTMSANESLTSQSFRFAEPFVGISASGSO-TAGTISINNAG 625
 DB 541 DEDFFVSRGGTTVNPNFRFLRTMNSGDELKYGFRRAFTTPTFTQIDIRTSIQGLSG 600
 QY 626 RQTFHFDFKIEFIPITATFEAEYDLERAQEAVALFTNTNPNRLKTDVTDYHIDQVNLVA 685
 DB 601 NGEVYIDKIEIIPVATATFEAEYDLERAQEAVALFTNTNPNRLKTDVTDYHIDQVNLVA 660
 QY 686 CLSDEFCLDEKRELEKVKYAKRLSDERNLLODPNFTSINKQDPFISTNEQSNFTSIHEQ 745
 DB 661 CLSDEFCLDEKRELEKVKYAKRLSDERNLLODPNFTSINKQDPFISTNEQSNFTSIHEQ 720
 QY 746 SEHGWWGSENIITQEGNDVFKENYVTLPGTFNECYPTLYQKIGESLKAITYQLRGYI 805
 DB 721 SEHGWWGSENIITQEGNDVFKENYVTLPGTFNECYPTLYQKIGESLKAITYQLRGYI 780
 QY 806 EDSQDLEIYLIRYNAKHETLDVPGTSEVWPLSVESPIGRCEPNRCAPHFENWPNLDSCS 865

Db 781 EDSQDLLEIYLIRNAKHETLDVPTESLWPLSVESPIGRCEPNRCAHPHFWPPDLDSC 840
QY 866 RDGEKCAHSHHSLDIDIGCTDLHENLGVVWVFKIKTQEGHARLGNLEFIEBKPLGEEA 925
Db 841 RDGEKCAHSHHSLDIDIGCTDLHENLGVVWVFKIKTQEGHARLGNLEFIEBKPLGEEA 900
QY 926 LSRVKRAEKWRDKREKLEQLQLETRKRVYTEAKVADALFVDSQYRNLQADTNIGMHAADKL 985
Db 901 LSRVKRAEKWRDKREKLEQLQLETRKRVYTEAKVADALFVDSQYRNLQADTNIGMHAADKL 960
QY 986 VHRIRAYLSLSVPCVNAEIEELEGRIITAIISLYDARNVKNKGDFNGLACVNVKXH 1045
Db 961 VHRIRAYLSLPIPCVNAEIEELEGRIITAIISLYDARNVKNKGDFNGLTCVNVKXH 1020
QY 1046 VDQQSHRSLVPIPEWAEVSQAVRVCPCGRGILRVYTAKEGEGECVTIHIENNTDE 1105
Db 1021 VDQQSHRSLVPIPEWAEVSQAVRVCPCGRGILRVYTAKEGEGECVTIHIENNTDE 1080
QY 1106 LKFNCEEEVYPTDTCNDYTAHQGTA---VCNSRNAGYEDAVEVDVTASVNVKPTY 1161
Db 1081 LKFNCEEEVYPTDTCNDYTAHQGTA---VCNSRNAGYEDAVEVDVTASVNVKPTY 1140
QY 1162 EESTYDVRDNCEYDRGVVNPPLPAGYMTKELEYFPETDKVWIEIGETEGKFIIVDSV 1221
Db 1141 EESTYDVRDNCEYDRGVVNPPLPAGYMTKELEYFPETDKVWIEIGETEGKFIIVDSV 1200
QY 1222 ELLMEE 1228
Db 1201 ELLMEE 1207

RESULT 7

US-08-459-595A-7

; Sequence 7, Application US/08459595A

; Patent No. 6018104

; GENERAL INFORMATION:

; APPLICANT: Koziel, Michael G.

; APPLICANT: Desai, Nalini M.

; APPLICANT: Lewis, Kelly S.

; APPLICANT: Kramer, Vance C.

; APPLICANT: Warren, Gregory W.

; APPLICANT: Evola, Stephen V.

; APPLICANT: Crossland, Lyle D.

; APPLICANT: Wright, Martha S.

; APPLICANT: Merlin, Ellis J.

; APPLICANT: Launis, Karen L.

; APPLICANT: Rothstein, Steven J.

; APPLICANT: Bowman, Cindy G.

; APPLICANT: Dawson, John L.

; APPLICANT: Dunder, Erik M.

; APPLICANT: Pace, Gary M.

; APPLICANT: Suttie, Janet L.

; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED

; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE

; NUMBER OF SEQUENCES: 94

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 6018104artis Corporation

; STREET: Patent & Trademark Dept., 520 White Plains

; STREET: Rd., POB 2005

; CITY: Tarrytown

; STATE: New York

; COUNTRY: USA

; ZIP: 10591-9005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/459,595A

; FILING DATE: 02-JUN-1995

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40403
; REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8582
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-459-595A-7

Query Match 87.4%; Score 5659.5; DB 3; Length 1207;

Best Local Similarity 89.4%; Pred. No. 0;

Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;

QY 27 MDLSPDARIEDSLCIAEGNNINPLVSASTVQTGINIAGRILGVLPFAGQIASFYSLV 86
Db 1 MDLLPDARIEDSLCIAEGNNIDPFVSASTVQTGINIAGRILGVLPFAGQIASFYSLV 60
QY 87 GELWPRGRDQWEIFLEHVEQLINQOITENARNALTALRQLGDSFRAYQOSLEDWLENRD 146
Db 61 GELWPRGRDQWEIFLEHVEQLINQOITENARNALTALRQLGDSFRAYQOSLEDWLENRD 120
QY 147 DARTSRVLYTQYIALELDFLNAMPFAIRNQEVPLLMVYAQAANLHLLLRDLASLFGSEF 206
Db 121 DARTSRVLYTQYIALELDFLNAMPFAIRNQEVPLLMVYAQAANLHLLLRDLASLFGSEF 180
QY 207 GLTSQBIQRYERQVQTRDYSDYCVIEWYNTGLNSLRGTNAASWVRVYQFRDLTLGVLD 266
Db 181 GLTSQBIQRYERQVQTRDYSDYCVIEWYNTGLNSLRGTNAASWVRVYQFRDLTLGVLD 240
QY 267 LVALPFSYDTRTYPINTSAQLTREVTDAIGATGVNMAWNNVNNAPSFAIETAVIRS 326
Db 241 LVALPFSYDTRTYPINTSAQLTREVTDAIGATGVNMAWNNVNNAPSFAIETAVIRS 300
QY 327 PHLLDFLEQLTIESTSRWSATRHMTYWRGHTTQSRPIGGGLNTSTHGSTNTSINPVRLS 386
Db 301 PHLLDFLEQLTIESTSRWSATRHMTYWRGHTTQSRPIGGGLNTSTHGSTNTSINPVRLS 360
QY 387 FFSRDVYVWTESYAGVLLWGIYLEPIHGVPVTRFNRPNQNTFERGTANYSQPYESPGQL 446
Db 361 FASRDVYVWTESYAGVLLWGIYLEPIHGVPVTRFNRPNQNTFERGTANYSQPYESPGQL 420
QY 447 KDSETELPETTERPNYESYSHRLSHGLISQSRVHVVPVSWTHRSADRNTTSSDITQ 506
Db 421 KDSETELPETTERPNYESYSHRLSHGLISQSRVHVVPVSWTHRSADRNTTSGPNRITQ 480
QY 507 IPLVKSFNLNSGTSVVGSGPGTGGDIIRTNVNGSVLSMGLNFNNTSLQRVVRVRYAASQ 566
Db 481 IPWKASELPQGTIVVRGPGTGGDIIRTNVNGSGFPIRVTVNGPLTORVIRGFYASTV 540
QY 567 TMLVRLVTGSGTTFDQGFPTMSANESLTSQSFRFAEPVVGISASGSQ-TAGISISNAG 625
Db 541 DFDFFVSRGGTIVNFRFLRTMNSGDELKYGNFVRAFTTPTFTQIQDIIRTSIQGLSG 600
QY 626 ROTFHEKDIEFIPITATFEAEYDLERAQEAVALFTNTNPRRLKTDVTDHIDQVSNLVA 685
Db 601 NGEVYIDKIEIIPVATATFEAEYDLERAQEAVALFTNTNPRRLKTDVTDHIDQVSNLVA 660
QY 686 CLSDEFLCKELLEKVKYAKELSDERNLLQDPNFTSINKOPDFISTNQSQNTSITHEQ 745
Db 661 CLSDEFLCKELLEKVKYAKELSDERNLLQDPNFTSINKOPDFISTNQSQNTSITHEQ 720


```

QY 746 SEHGWSGSENIITQEGNDVFNKENVVILPGTFNECYPYLYQKIGESLKYATRYQLRGYI 805
Db 721 SEHGWSGSENIITQEGNDVFNKENVVILPGTFNECYPYLYQKIGESLKYATRYQLRGYI 780
QY 806 EDSQDLBIYLIRYNAXHETLDVPGTESVWPLSVESPIGRCEPNRCAPHFEMWPDLDSC 865
Db 781 EDSQDLBIYLIRYNAXHETLDVPGTESVWPLSVESPIGRCEPNRCAPHFEMWPDLDSC 840
QY 866 RDEKCAHSHHSLDIDIGCTDLHENLGVVWVFKIKTQGHARLGNLEFIEBKPLLGEA 925
Db 841 RDEKCAHSHHSLDIDIGCTDLHENLGVVWVFKIKTQGHARLGNLEFIEBKPLLGEA 900
QY 926 LSRVKRAEKWRDKREKLETKRVYTEAEAVDALFVDSQYNRLOADTNIGMHAADKL 985
Db 901 LSRVKRAEKWRDKREKLETKRVYTEAEAVDALFVDSQYNRLOADTNIGMHAADKL 960
QY 986 VHRIRAYLSLSVPGVNAEIPFEELBGRITITAIISLYDARNVVKNGDFNNGLACWNVKGH 1045
Db 961 VHRIRAYLSLSVPGVNAEIPFEELBGRITITAIISLYDARNVVKNGDFNNGLACWNVKGH 1020
QY 1046 VDVQOQSHRSVLVPEWEAEVSQAVRVCGRGVLRTAYKEGEGECVTIHEIENNTDE 1105
Db 1021 VDVQOQSHRSVLVPEWEAEVSQAVRVCGRGVLRTAYKEGEGECVTIHEIENNTDE 1080
QY 1106 LKFKNCEEEVYPTDGTCTNDYTAHQGTAAA--VCNSRNAGYEDAYEVDVTTASVNVKPTY 1161
Db 1081 LKFKNCEEEVYPTDGTCTNDYTAHQGTAAA--VCNSRNAGYEDAYEVDVTTASVNVKPTY 1140
QY 1162 EETFTDVRDNHCEYDRGVNYPPLPAGVMTKELEYFPFTDKWIBIGTEGKFIVDSV 1221
Db 1141 EETFTDVRDNHCEYDRGVNYPPLPAGVMTKELEYFPFTDKWIBIGTEGKFIVDSV 1200
QY 1222 ELLIMEE 1228
Db 1201 ELLIMEE 1207

```

RESULT 8

```

US-08-459-504B-7
; Sequence 7, Application US/08459504B
; Patent No. 6075185
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 6075185artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,504B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/459,595
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1577/CIP/DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-459-504B-7

Query Match 87.4%; Score 5659.5; DB 3; Length 1207;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;

QY 27 MDLSPDARIEDSLCIAEGNNINPLVSASTVQTGINIAGRIILGVLPFAGQIASFYSLV 86
Db 1 MDLSPDARIEDSLCIAEGNNIDPVSASTVQTGINIAGRIILGVLPFAGQIASFYSLV 60
QY 87 GELWPRGRDQWEIFLEHVEQLINQITENARNALTALRQLGLDPSFRAYQOSLEDWLENRD 146
Db 61 GELWPRGRDQWEIFLEHVEQLINQITENARNALTALRQLGLDPSFRAYQOSLEDWLENRD 120
QY 147 DARTSRVLYTQYIALDELDFINAMPLFAIRNOEVPILMVTAQAANLHLLLRDASLFGSEF 206
Db 121 DARTSRVLYTQYIALDELDFINAMPLFAIRNOEVPILMVTAQAANLHLLLRDASLFGSEF 180
QY 207 GLTSQEIQRYERQVETRDYSDYCVWYNTGLNSLRGTNAASWRYNQPFRRLDTLGLVD 266
Db 181 GLTSQEIQRYERQVETRDYSDYCVWYNTGLNSLRGTNAASWRYNQPFRRLDTLGLVD 240
QY 267 LVALFPSYDTRTPINTSAQLTREVTDAIGATGVNMAWNNWNNNAPSFAIETAVIRS 326
Db 241 LVALFPSYDTRTPINTSAQLTREVTDAIGATGVNMAWNNWNNNAPSFAIETAVIRS 300
QY 327 PHLLDFLEQLTIFSTSRWSATHMTYWRGHTIQSRPIGGGLNTSTHGSTNTSINPVLRS 386
Db 301 PHLLDFLEQLTIFSTSRWSATHMTYWRGHTIQSRPIGGGLNTSTHGSTNTSINPVLRS 360
QY 387 FFSRDVVTESVAGVLLWGLYLEPIHGVTVPFRNPNPONTFERGTANTSQPYESPGQLQ 446
Db 361 FASRDVVTESVAGVLLWGLYLEPIHGVTVPFRNPNPONTFERGTANTSQPYESPGQLQ 420
QY 447 KQSETELPPTTERPNYESHRLSHIGLSQSRVHPVYVSWTHRSADRNTNTSSDITQ 506
Db 421 KQSETELPPTTERPNYESHRLSHIGLSQSRVHPVYVSWTHRSADRNTNTSSDITQ 480
QY 507 IPLVKSFNLSGTSVVGPGFTGGDIIRTNVNGSVLSMGLNFNNTSLQVRVVRVAASQ 566
Db 481 IPVKSASELPQGTVVVRGPGFTGGDIIRTNVNGSVLSMGLNFNNTSLQVRVVRVAASQ 540
QY 567 TMVLRTVVGSTTFDQGFPTSMANESLTSQSFRFAEFVPGISAGSQ-TAGISISNAG 625
Db 541 DFDFFVSRGGTTVNNFRFLRTMNSGDELKXGVFVRAFTPTFTQIQDIIRTSIQLSG 600
QY 626 RQTFHFDKIEFIPITATFAEYDLERAQBAVNALFNTNPRRLKTDVTDYHIDQVNLVA 685

```

Db 601 NGEVYIDKIEIIPVTATFEAEYDLERAQEAVALFTNTNPRRLKTDVTDYHIDQVSNLVA 660
Qy 686 CLSDEFCLDEKRELLEKVKYAKRLSDERNLLQDNFTSINKQDPFISTNEQSNFTSIHEQ 745
Db 661 CLSDEFCLDEKRELLEKVKYAKRLSDERNLLQDNFTSINKQDPFISTNEQSNFTSIHEQ 720
Qy 746 SEHGWGSENITIOEGNDVFKENVTLPCTNECYPTLYQKIGESLKYATRYQLRGYI 805
Db 721 SEHGWGSENITIOEGNDVFKENVTLPCTNECYPTLYQKIGESLKYATRYQLRGYI 780
Qy 806 EDSQDLIYLIRYNAKHETLDVPGTESVWPLSVESPIGRCEPNRCAPHFENWPDLCSC 865
Db 781 EDSQDLIYLIRYNAKHETLDVPGTESVWPLSVESPIGRCEPNRCAPHFENWPDLCSC 840
Qy 866 RDGEKCAHSHHFLSDIDIGCTDLHENLGVWVVPFKIKTQEGHARLGNLEFIEEKPLGEEA 925
Db 841 RDGEKCAHSHHFLSDIDIGCTDLHENLGVWVVPFKIKTQEGHARLGNLEFIEEKPLGEEA 900
Qy 926 LSRVKRAEKKWRDKREKLOLETKREYVTEAKEAVDALFVDSQYNRLQADTWIGMHAADKL 985
Db 901 LSRVKRAEKKWRDKREKLOLETKREYVTEAKEAVDALFVDSQYNRLQADTWIGMHAADKL 960
Qy 986 VHRIRZAYLSLSVPGVNAEIPFEELEGRITITSLYDARNVVKNGDFNNGLACWNVKGH 1045
Db 961 VHRIRZAYLSLSVPGVNAEIPFEELEGRITITSLYDARNVVKNGDFNNGLACWNVKGH 1020
Qy 1046 VDVQSHRSVLVTPWEAEVQSVQAVRCPGCGYILRTVAYKEGEGCVTIHEIENNTDE 1105
Db 1021 VDVQSHRSVLVTPWEAEVQSVQAVRCPGCGYILRTVAYKEGEGCVTIHEIENNTDE 1080
Qy 1106 LKPKNCEEEVYPTDTCNDYTAHOCTA----VCNSRNAGYEDAYEVDTTASVNYKPTY 1161
Db 1081 LKPKNCEEEVYPTDTCNDYTAHOCTA----VCNSRNAGYEDAYEVDTTASVNYKPTY 1140
Qy 1162 EEEYTDVRDNDHCEYDRGVYVNPPLPAGYMTKELEFPFETDKWVIBIGTEGKFIVDVS 1221
Db 1141 EEEYTDVRDNDHCEYDRGVYVNPPLPAGYMTKELEFPFETDKWVIBIGTEGKFIVDVS 1200
Qy 1222 ELLMEE 1228
Db 1201 ELLMEE 1207

RESULT 9

US-08-459-444-7
; Sequence 7, Application US/08459444A
; Patent No. 6121014
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; Desai, Nalini M.
; Lewis, Kelly S.
; Kramer, Vance C.
; Warren, Gregory W.
; Evola, Stephen V.
; Crossland, Lytle D.
; Wright, Martha S.
; Merlin, Ellis J.
; Launis, Karen L.
; TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED
; NUCLEIC ACID CODING SEQUENCE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,444A
FILING DATE: 02-Jun-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-18805/P1/CGC1577/CIP/DIV6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1207 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-459-444-7

Query Match 87.4%; Score 5659.5; DB 3; Length 1207;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;

Qy 27 MDLSPDARIEDSLCIAEGNNINPLVSAVSTVQTGINIAGRILGVLPVFPAGQIASFYSLV 86
Db 1 MDLSPDARIEDSLCIAEGNNIDPFVSAVSTVQTGINIAGRILGVLPVFPAGQIASFYSLV 60
Qy 87 GELWPRGRDQWEIFLEHVEQLINQOITENARNALALQGLGDSFRAYQOSLEDWLENRD 146
Db 61 GELWPRGRDQWEIFLEHVEQLINQOITENARNALALQGLGDSFRAYQOSLEDWLENRD 120
Qy 147 DARTSRVLYTQYIALELDLFLNAMPFAIRNQEVPLLMVYAQAANHLHLLLRDASLFQSEF 206
Db 121 DARTSRVLYTQYIALELDLFLNAMPFAIRNQEVPLLMVYAQAANHLHLLLRDASLFQSEF 180
Qy 207 GLTSQEIQRVYRQVEQTRDYSYCVWEYNTGLNSLGTNAASWVRVYQNRDRDLTLGVLD 266
Db 181 GLTSQEIQRVYRQVEQTRDYSYCVWEYNTGLNSLGTNAASWVRVYQNRDRDLTLGVLD 240
Qy 267 LVALFPSYDTRTYPINTSAQLTREVTDAIGATGVNMAAMNWNNAAPSAIAETAVIRS 326
Db 241 LVALFPSYDTRTYPINTSAQLTREVTDAIGATGVNMAAMNWNNAAPSAIAETAVIRS 300
Qy 327 PHLLDFLEQLTIFTSGRWSATRMVYRGTIQSRPIGGGLNTSTHGNTNTSINPVTLR 386
Db 301 PHLLDFLEQLTIFTSGRWSATRMVYRGTIQSRPIGGGLNTSTHGNTNTSINPVTLR 360
Qy 387 FFSRDVYWTESYAGVLLWGIYLEPIHCVPTVRNFRNPONTFERGTANYQPYESRGLQ 446
Db 361 FASRDVYWTESYAGVLLWGIYLEPIHCVPTVRNFRNPONTFERGTANYQPYESRGLQ 420
Qy 447 KDSETELPPTTERPNYESYSHRLSHIGLSQSRVHVVPVSWTHRSADRNTISSDSITQ 506
Db 421 KDSETELPPTTERPNYESYSHRLSHIGLSQSRVHVVPVSWTHRSADRNTIGPNRITO 480
Qy 507 IPLVKSPNLNSGTVSGPGFTGDIIRTNVNGSVLSMGLNFNNTSLQRVVRVRYAASQ 566
Db 481 IPWKASSELPGQITVVRGPGFTGDIIRTNVNGSVLSMGLNFNNTSLQRVVRVRYAASQ 540
Qy 567 TMVLRTVVGSTTDFQGFPPSTMSANESLTSQSRFAEFPPVGISASGQ-TAGISISNAG 625
Db 541 DFFDFVSRGGTTVNFRFLETMSGDELKGNFVRRAFTPTFTTQIDIRTSIQGLSG 600
Qy 626 RQTFHFDEKIEFIPITATFEAEYDLERAQEAVALFTNTNPRRLKTDVTDYHIDQVSNLVA 685
Db 601 NGEVYIDKIEIIPVTATFEAEYDLERAQEAVALFTNTNPRRLKTDVTDYHIDQVSNLVA 660

686 CLSDFCLDEKRELEKVKYAKRLSDERNLLQDPNFTSINKQDPFISTNEQSNFTSIHQ 745
661 CLSDFCLDEKRELEKVKYAKRLSDERNLLQDPNFTSINKQDPFISTNEQSNFTSIHQ 720
746 SEHGWSGSENIITQEGNDVFKENYVTLPGTFNECYPTLYQKIGESLKYATRYQLRGI 805
721 SEHGWSGSENIITQEGNDVFKENYVTLPGTFNECYPTLYQKIGESLKYATRYQLRGI 780
806 EDSQDLIYLIRYNKAKHETLDVPGTESVWPLSVESPIGRCEPNRCAPHFENPDLCSC 865
781 EDSQDLIYLIRYNKAKHETLDVPGTESVWPLSVESPIGRCEPNRCAPHFENPDLCSC 840
866 RDGEKCAHSHHFSLDIDIGCTDLHENLGVWVFKIKTQEGHARLGNLEFIEEKPILGEA 925
841 RDGEKCAHSHHFSLDIDVGCTDLHENLGVWVFKIKTQEGHARLGNLEFIEEKPILGEA 900
926 LSRVKRAEKKWRDKREKLQLETKRVYTEAKEAVDALFVDSQYNRQLQADTNGIMHAADKL 985
901 LSRVKRAEKKWRDKREKLQLETKRVYTEAKEAVDALFVDSQYNRQLQADTNGIMHAADKL 960
986 VHRIRAYLSLSVTPGNAEIPFELSGRIITAIISLYDARNVKNQGFNNGLTCAWVYKGH 1045
961 VHRIRAYLSVTPGNAEIPFELSGRIITAIISLYDARNVKNQGFNNGLTCAWVYKGH 1020
1046 VDVOQSHRSVLVIPWEAEVSQAVRVCPCRGYILRVTAKEGEGECVTHIEIENNTDE 1105
1021 VDVOQSHRSVLVIPWEAEVSQAVRVCPCRGYILRVTAKEGEGECVTHIEIENNTDE 1080
1106 LKFNCEBEEVYPTDTCNDYTAHQCTA ---VCNSRNAGYEDAYEVDVTTASVNYKPTY 1161
1081 LKFNCEBEEVYPTDTCNDYTAHQCTA ---VCNSRNAGYEDAYEVDVTTASVNYKPTY 1140
1162 EETTYDVRDNHCEYDRGVNPPPLPAGVMTKELEYFPEPTDKWIEIGTEGKFIVDSV 1221
1141 EETTYDVRDNHCEYDRGVNPPVPGVYVTKLEYFPEPTDKWIEIGTEGKFIVDSV 1200
1222 ELLMEE 1228
1201 ELLMEE 1207

RESULT 10

US-09-053-549-8
Sequence 8, Application US/09053549
Patent No. 6121521
GENERAL INFORMATION:
APPLICANT: Desai, Nalini
TITLE OF INVENTION: No. 6121521el Insecticidal Protein and Gene
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6121521artis Corporation
STREET: 3054 Cornwallis Rd.
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/053,549
FILING DATE: 01-APR-1998
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: 40,403
REFERENCE/DOCKET NUMBER: CGC 1995
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8582
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 1207 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-053-549-8

Query Match. 87.4%; Score 5659.5; DB 3; Length 1207;

Best Local Similarity 89.4%; Pred. No. 0;

Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;

QY 27 MDLSPDARIEDSUCIAEGNNINPLVSASTVQTGINIAGRILGVLPFAQGIASFYSFLV 86
DB 1 MDLSPDARIEDSUCIAEGNNIDPFVSASTVQTGINIAGRILGVLPFAQGIASFYSFLV 60
QY 87 GELWPRGRDQWEIFLEHVEQLINQOITENARNALTALRQGLGDSFRAYQOQSLDWLENRD 146
DB 61 GELWPRGRDQWEIFLEHVEQLINQOITENARNALTALRQGLGDSFRAYQOQSLDWLENRD 120
QY 147 DARTSRVLTQYIALDELDFLNAMPLFAIRNOEVPILMVYAQAANILHLLLRDASLFGSEF 206
DB 121 DARTSRVLTQYIALDELDFLNAMPLFAIRNOEVPILMVYAQAANILHLLLRDASLFGSEF 180
QY 207 GLTSQIQRYYERQVROTSDYSCYEWYNTGLNSLRGTNAASWRYNFRRLDITLGVL 266
DB 181 GLTSQIQRYYERQVROTSDYSCYEWYNTGLNSLRGTNAASWRYNFRRLDITLGVL 240
QY 267 LVALPFSYDTRTPINTSAQLTREVTDAIGATGVNMAWNNYNNAPSAIETAVIRS 326
DB 241 LVALPFSYDTRTPINTSAQLTREVTDAIGATGVNMAWNNYNNAPSAIETAVIRS 300
QY 327 PHLLDFLEOLTIFSTSSRSATRHMTYWRGHTIQSRPIGGGLNTSTHGTNTSINPVRLS 386
DB 301 PHLLDFLEOLTIFSTSSRSATRHMTYWRGHTIQSRPIGGGLNTSTHGTNTSINPVRLS 360
QY 387 FFSRDVYVYESVAGVLLWGIYLEPIHGVPTVFRNPNQNTFERGTANTSQPYESPLQL 446
DB 361 FASRDVYVYESVAGVLLWGIYLEPIHGVPTVFRNPNQNTFERGTANTSQPYESPLQL 420
QY 447 KQSETELPETTERPNYESYSHRLSHIGLISQSRVHVVPVYSWTHRSADRTNTISSSIQ 506
DB 421 KQSETELPETTERPNYESYSHRLSHIGLISQSRVHVVPVYSWTHRSADRTNTIGPNRIQ 480
QY 507 IPLVKSFNLSGTSVVSQPGFTGGDIIRTNVNGSVLSMGLNFNNTSLQRYRVRVRAAQ 566
DB 481 IPWKRASELPQGTIVVRGPGFTGGDIIRTNVNGSVLSMGLNFNNTSLQRYRVRVRAAQ 540
QY 567 TWVLRVTVGGSTTFDQGFPTMSANESLTSQSFRFAEFVVGISASGSQ-TAGISISNNAG 625
DB 541 DPDFVVSRGTTVNNFRFLRTWNSGDELKYGVRRAFTTPFTTQIDIIRTSIQGLSG 600
QY 626 RQTFHDKLEFIPITATPEAYDLERAQAVNALFTNTNPRRLKTDVTYHIDQVNLVA 685
DB 601 NGEVYIDKIEIIPVATPEAYDLERAQAVNALFTNTNPRRLKTDVTYHIDQVNLVA 660
QY 686 CLSDFCLDEKRELEKVKYAKRLSDERNLLQDPNFTSINKQDPFISTNEQSNFTSIHQ 745
DB 661 CLSDFCLDEKRELEKVKYAKRLSDERNLLQDPNFTSINKQDPFISTNEQSNFTSIHQ 720
QY 746 SEHGWSGSENIITQEGNDVFKENYVTLPGTFNECYPTLYQKIGESLKYATRYQLRGI 805
DB 721 SEHGWSGSENIITQEGNDVFKENYVTLPGTFNECYPTLYQKIGESLKYATRYQLRGI 780
QY 806 EDSQDLIYLIRYNKAKHETLDVPGTESVWPLSVESPIGRCEPNRCAPHFENPDLCSC 865
DB 781 EDSQDLIYLIRYNKAKHETLDVPGTESVWPLSVESPIGRCEPNRCAPHFENPDLCSC 840
QY 866 RDGEKCAHSHHFSLDIDIGCTDLHENLGVWVFKIKTQEGHARLGNLEFIEEKPILGEA 925
DB 841 RDGEKCAHSHHFSLDIDVGCTDLHENLGVWVFKIKTQEGHARLGNLEFIEEKPILGEA 900
QY 926 LSRVKRAEKKWRDKREKLQLETKRVYTEAKEAVDALFVDSQYNRQLQADTNGIMHAADKL 985

Db 901 LSRVRAEKKRDKREKLEQLQLETKVYTEAKEAVDALFVDSQYDELQADTNIGMHAADKL 960
 Qy 986 VHRIREAYLSVLPVGNNAEIPFEELEGRITITSLYDARNVVKNGDFNGLACWNVKGH 1045
 Db 961 VHRIREAYLSVLPVGNNAEIPFEELEGRITITSLYDARNVVKNGDFNGLTCWNVKGH 1020
 Qy 1046 VDVOQHRSVLPVPEAEVSQARVCPGCGYILRVYAYKEGVCVTHIEHNNTDE 1105
 Db 1021 VDVOQHRSVLPVPEAEVSQARVCPGCGYILRVYAYKEGVCVTHIEHNNTDE 1080
 Qy 1106 LKFKNEEEVYPTDTCNDYTAHQGTA---VCNSRNAGYEDAYEVDVTTASVNYKPTY 1161
 Db 1081 LKFKNEEEVYPTDTCNDYTAHQGTAAGACADACNSNAGYEDAYEVDVTTASVNYKPTY 1140
 Qy 1162 EETTYDVRDNCEYDRGVNYPPLPAGYNTKELEYFPETDKWIBIGTEGKFI VDSV 1221
 Db 1141 EETTYDVRDNCEYDRGVNYPPLPAGYNTKELEYFPETDKWIBIGTEGKFI VDSV 1200
 Qy 1222 ELLMEE 1228
 Db 1201 ELLMEE 1207

RESULT 11

US-09-547-422-7
 ; Sequence 7, Application US/09547422
 ; Patent No. 6320100
 ; GENERAL INFORMATION:
 ; APPLICANT: Koziel, Michael G.
 ; Desai, Nalini M.
 ; Lewis, Kelly S.
 ; Kramer, Vance C.
 ; Warren, Gregory W.
 ; Evola, Stephen V.
 ; Crossland, Lyle D.
 ; Wright, Martha S.
 ; Merlin, Ellis J.
 ; Lanis, Karen L.
 ; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
 ; INSECTICIDAL ACTIVITY IN MAIZE
 ; NUMBER OF SEQUENCES: 94
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 6320100artis Agribusiness Biotechnology Research, Inc.
 ; STREET: 3054 Cornwallis Road
 ; CITY: Research Triangle Park
 ; STATE: NC
 ; COUNTRY: USA
 ; ZIP: 27709
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/547,422
 ; FILING DATE: 11-Apr-2000
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/459,595
 ; FILING DATE: 02-JUN-1995
 ; APPLICATION NUMBER: US 07/951,715
 ; FILING DATE: 25-SEP-1992
 ; APPLICATION NUMBER: US 07/772,027
 ; FILING DATE: 04-OCT-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meigs, J. Timothy
 ; REGISTRATION NUMBER: 38,241
 ; REFERENCE/DOCKET NUMBER: S-18805H
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (919) 541-8587
 ; TELEFAX: (919) 541-8689
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:

; LENGTH: 1207 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
 US-09-547-422-7
 Query Match 87.4%; Score 5659.5; DB 3; Length 1207;
 Best Local Similarity 89.4%; Pred. No. 0;
 Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;
 Qy 27 MDLSPDARIEDSICIAEGNNINPLVSASTVQTGINTAGRIILGVLPFAGQIASFYFLV 86
 Db 1 MDLSPDARIEDSICIAEGNNIDPFVSASTVQTGINTAGRIILGVLPFAGQIASFYFLV 60
 Qy 87 GELWPRGRDQWEIFLEHVEQLINQOITENARNALARLOGLGDSFRAYQOSLEDWLNRD 146
 Db 61 GELWPRGRDQWEIFLEHVEQLINQOITENARNALARLOGLGDSFRAYQOSLEDWLNRD 120
 Qy 147 DARTSRVLYTQYIALELDLFLNAMPLFAIRNQEVPLLMVYAQAANLHLLLRDASLFGSEF 206
 Db 121 DARTSRVLYTQYIALELDLFLNAMPLFAIRNQEVPLLMVYAQAANLHLLLRDASLFGSEF 180
 Qy 207 GLTSQEIORYEROVQTRDYSYCVWYNTGLNSLRGTNAASWVRVNOFRDRLTGLVD 266
 Db 181 GLTSQEIORYEROVQTRDYSYCVWYNTGLNSLRGTNAASWVRVNOFRDRLTGLVD 240
 Qy 267 LVALFPSYDTRTPINTSAQLTREVTDAIGATGVNMAWNNYNNAPSAIETAIVRS 326
 Db 241 LVALFPSYDTRTPINTSAQLTREVTDAIGATGVNMAWNNYNNAPSAIETAIVRS 300
 Qy 327 PHLLDFLEQLTIFSTSRMSATRMHYWRGHTIQSRPIGGGLNTSTHGSTNTSINPVRLS 386
 Db 301 PHLLDFLEQLTIFSASSRWSNTRHMTYWRGHTIQSRPIGGGLNTSTHGSTNTSINPVTLR 360
 Qy 387 FFSRDVYTESYAGVLLWGIYLEPIHGVPVTRFNPNQPTFERGTANYSQPYESPGLQ 446
 Db 361 FASRDVYTESYAGVLLWGIYLEPIHGVPVTRFNPNQPTFERGTANYSQPYESPGLQ 420
 Qy 447 KQSETELPETTERPNVYESYSHRLSHIGLISQSRVHVVPVSWTHRSADRNTNTSSDSIT 506
 Db 421 KQSETELPETTERPNVYESYSHRLSHIGLISQSRVHVVPVSWTHRSADRNTNTCPNRIQ 480
 Qy 507 IPLVKFNLSNGTSVSGPGFTGDIIRTNVNGSVLSMGLNFNNTSLQRVVRVRYAASQ 566
 Db 481 IPWKASLPQGTTVVRGPGFTGDIIRTNVNGSVLSMGLNFNNTSLQRVVRVRYAASQ 540
 Qy 567 TMVLRYTVGSGTTFDQGFSTMSANESLTSQSRFAEPFVGISASCSQ-TAGISISNAG 625
 Db 541 DFDFVSRGTTVNNFRFLRTMNSGDELKYGNVFRAFTPTFTTQTDIIRTSIQGLSG 600
 Qy 626 RQTFHFDKIEFIPITATFAEYDLERAQEAVALFTNTNPRRLKTDVTDYHIDOVSNLVA 685
 Db 601 NGEVYIDKIEIIPVATFAEYDLERAQEAVALFTNTNPRRLKTDVTDYHIDOVSNLVA 660
 Qy 686 CLSDEFCLDEKRELLEKVKYAKELSDERNLLOPNFTSINKQPDFISTNQSNFTSHEQ 745
 Db 661 CLSDEFCLDEKRELLEKVKYAKELSDERNLLOPNFTSINKQPDFISTNQSNFTSHEQ 720
 Qy 746 SEHGWNGSENITIOEGNDVFKENYVTLPGTFNFCYPTLYLQKIGESLKYATRYQLRGI 805
 Db 721 SEHGWNGSENITIOEGNDVFKENYVTLPGTFNFCYPTLYLQKIGESLKYATRYQLRGI 780
 Qy 806 EDSQDLEIYLIRYNAKHETLDVPGTESVMPLESPIGRGEPNRCAPHFEMWPDLDCCS 865
 Db 781 EDSQDLEIYLIRYNAKHETLDVPGTESVMPLESPIGRGEPNRCAPHFEMWPDLDCCS 840
 Qy 866 RDGEKCAHSHHPSLDIDIGCTDLHENLGVVYFKIKTQEGHARLGNLEFIEEKLGEA 925
 Db 841 RDGEKCAHSHHPSLDIDIGCTDLHENLGVVYFKIKTQEGHARLGNLEFIEEKLGEA 900
 Qy 926 LSRVKAERKWRDKREKLEQLQLETKVYTEAKEAVDALFVDSQYDELQADTNIGMHAADKL 985

Db 901 LSRVRAEKKWRKKEQLQLETKRVTYEAKEAVDALFVDSQYDRLQADTNIGMIHAADKL 960
Qy 986 VHRITREAYLSVLPVGNAAEIPFELSGRIITAIISLYDARNVVKNGDPNGLACWNKGH 1045
Db 961 VHRITREAYLSVLPVGNAAEIPFELSGRIITAIISLYDARNVVKNGDPNGLACWNKGH 1020
Qy 1046 VDVOQSHHRSVLVPEWAEVSQAVRCPGRGYILRTVAYKEGEGCVTIHETENNTE 1105
Db 1021 VDVOQSHHRSVLVPEWAEVSQAVRCPGRGYILRTVAYKEGEGCVTIHETENNTE 1080
Qy 1106 LKFKNCBEEVYPTDTGNDYTAHQCTA-----VCNSRNAGYEDAYEVDVTTASVNYKPT 1161
Db 1081 LKFKNCBEEVYPTDTGNDYTAHQCTAGCADACNSRNAGYEDAYEVDVTTASVNYKPT 1140
Qy 1162 EEEYTTDVRDNHCEYDRGVNYPPLPAGYMTKELEYFPETDKWIBIGTEGKFIVDSV 1221
Db 1141 EEEYTTDVRDNHCEYDRGVNYPVPAGYMTKELEYFPETDKWIBIGTEGKFIVDSV 1200
Qy 1222 ELLLMEE 1228
Db 1201 ELLLMEE 1207

RESULT 12

US-09-988-462-7
; Sequence 7, Application US/09988462
; Patent No. 6720488
; GENERAL INFORMATION:
; APPLICANT: Kozziel, Michael G.
; Desai, Nalini M.
; Lewis, Kelly S.
; Kramer, Vance C.
; Warren, Gregory W.
; Evola, Stephen V.
; Crossland, Lydie D.
; Wright, Martha S.
; Merlin, Ellis J.
; Launis, Karen L.

TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
INSECTICIDAL ACTIVITY IN MAIZE

NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: Syngenta Biotechnology, Inc.
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/988,462
FILING DATE: 20-No. 6720488-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/547,422
FILING DATE: 11-APR-2000
APPLICATION NUMBER: US 08/459,504
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: US 07/951,715
FILING DATE: 25-SEP-1992
APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-188051
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689

; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1207 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-988-462-7

Query Match 87.4%; Score 5659.5; DB 4; Length 1207;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;

Qy 27 MDLSPARIEDSLCIAEGNNINPLVSASTVTCINAGTILGVLGVFPFAGQIASFSFLV 86
Db 1 MDLLPARIEDSLCIAEGNNIDPFVSASTVTCINAGTILGVLGVFPFAGQIASFSFLV 60

Qy 87 GELWPRGRDQWEIFLEHVEQLINQOITENARTALARLOGLGDSFRAYQOSLEDWLENRD 146
Db 61 GELWPRGRDQWEIFLEHVEQLINQOITENARTALARLOGLGDSFRAYQOSLEDWLENRD 120

Qy 147 DARTRSVLYTQYIALBELDFLNAMPLFAIRNQEVPLLMVYAQAANLHLLLRDASLFGSEF 206
Db 121 DARTRSVLYTQYIALBELDFLNAMPLFAIRNQEVPLLMVYAQAANLHLLLRDASLFGSEF 180

Qy 207 GLTSQEIQRYYEROVETRDYSDYCVWEYNTGLNSLRGTNAASWVRYNQRFRDLTIGVLD 266
Db 181 GLTSQEIQRYYEROVETRDYSDYCVWEYNTGLNSLRGTNAASWVRYNQRFRDLTIGVLD 240

Qy 267 LVALPFSYDTRTYPIINTSAQLTREVTDAIGATGVNMAWMNNNAPSFALETAVIRS 326
Db 241 LVALPFSYDTRTYPIINTSAQLTREVTDAIGATGVNMAWMNNNAPSFALETAVIRS 300

Qy 327 PHLLDFLEQLTTFSTSSRWASATRHMTYWRGHTIQSRPICGGLNTSHGTSNTSINPVRLS 386
Db 301 PHLLDFLEQLTTFSTSSRWASATRHMTYWRGHTIQSRPICGGLNTSHGTSNTSINPVRLS 360

Qy 387 FFSRDVVTESYAGVLLWGIYLEPIHGVPTVFNFNPNONTFERGTANYSQPVESPLQL 446
Db 361 FFSRDVVTESYAGVLLWGIYLEPIHGVPTVFNFNPNONTFERGTANYSQPVESPLQL 420

Qy 447 KDSETELPETTERPNYESYSHRLSHIGLSRHHVVPVYVSWTHRSADRTNTSSDITQ 506
Db 421 KDSETELPETTERPNYESYSHRLSHIGLSRHHVVPVYVSWTHRSADRTNTSSDITQ 480

Qy 507 IPLVKSFNLSGTSVVGSGFTGGDIIRTNVNGSVLSMGLNFNNTSLQRYRVRVRAASQ 566
Db 481 IPLVKSFNLSGTSVVGSGFTGGDIIRTNVNGSVLSMGLNFNNTSLQRYRVRVRAASQ 540

Qy 567 TWVLRVTVGSTTFDQGPSTMSANESLTSQSFRAEPFVPGISASGSQ-TAGISISNAG 625
Db 541 TWVLRVTVGSTTFDQGPSTMSANESLTSQSFRAEPFVPGISASGSQ-TAGISISNAG 600

Qy 626 RQTFHFDKIEFIPITATPEAYDLERAQEAVALFTNTNPRKLKTDVTDYHIDQVSNLVA 685
Db 601 RQTFHFDKIEFIPITATPEAYDLERAQEAVALFTNTNPRKLKTDVTDYHIDQVSNLVA 660

Qy 686 CLSDEPCLDEKELLEKVKYAKRLSDERNLIODPNFTSINKQPDFISTNEQSNFTSIHQ 745
Db 661 CLSDEPCLDEKELLEKVKYAKRLSDERNLIODPNFTSINKQPDFISTNEQSNFTSIHQ 720

Qy 746 SEHGWMGSENITIQENDVFKENYVTLPGTFNECYPTYLYQKIGSELKAYTRYQLRGYI 805
Db 721 SEHGWMGSENITIQENDVFKENYVTLPGTFNECYPTYLYQKIGSELKAYTRYQLRGYI 780

Qy 806 EDSQDLIELIIRYNKAKHETLDVPGTESVWPLSVESPIGRCGEPNRCAPHEFNPDLDSCS 865
Db 781 EDSQDLIELIIRYNKAKHETLDVPGTESVWPLSVESPIGRCGEPNRCAPHEFNPDLDSCS 840

Qy 866 RDGEKCAHSHHFSLDIDIGCTDLHENLGVWVFKIKTOEGHARLGNLFIEKPLLGEA 925
Db 841 RDGEKCAHSHHFSLDIDIGCTDLHENLGVWVFKIKTOEGHARLGNLFIEKPLLGEA 900

926	Qy	LSRVGAEKWDRKREKJQLETRKRVYTEAKEAVDALFVDSQYNRLQADTNIGMHAADKL	985
901	Db	LSRVGAEKWDRKREKJQLETRKRVYTEAKEAVDALFVDSQYDRLOADTNIGMHAADKL	960
986	Qy	VHRIREAYLSELVIFGVNAEIPFEELEGRITAITLSLYDAENVVKNGDFNNGLACWNVKGH	1045
961	Db	VHRIREAYLSELVIFGVNAEIPFEELEGHITAITLSLYDAENVVKNGDFNNGLTCWNVKGH	1020
1046	Qy	VDVQOQSHRSVLVIPWEAEVSQARVCPGCGVILRVYAKEGVGECVTHIEHNNTDE	1105
1021	Db	VDVQOQSHRSDLVIPWEAEVSQARVCPGCGVILRVYAKEGVGECVTHIEHNNTDE	1080
1106	Qy	LKFKNCEEEVYPTDGTCDNYTAHQGTA---VCNRSNAGYEDAEVDVTASVNYKPTY	1161
1081	Db	LKFKNREEEVYPTDGTCDNYTAHQGTACACACNRSNAGYEDAEVDVTASVNYKPTY	1140
1162	Qy	EEETTYDVRDNHCEYDRGVNYPPLPAGMYKTELEYFPETDKWIEIGTEGKFIVDSV	1221
1141	Db	EEETTYDVRDNHCEYDRGVNYPVPVAGVYKTELEYFPETDTVWIEIGTEGKFIVDSV	1200
1222	Qy	ELLIMEE 1228	
1201	Db	ELLIMEE 1207	

RESULT 13

```

US-09-053-549-2
; Sequence 2, Application US/09053549
; Patent No. 6121521
; GENERAL INFORMATION:
; APPLICANT: Desai, Nalini
; TITLE OF INVENTION: No. 6121521el Insecticidal Protein and Gene
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6121521artis Corporation
; STREET: 3054 Cornwallis Rd.
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,549
; FILING DATE: 01-APR-1998
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: 40,403
; REFERENCE/DOCKET NUMBER: CGC 1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1227 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-053-549-2

```

```

Query Match      83.9%; Score 5436.5; DB 3; Length 1227;
Best Local Similarity 84.3%; Pred. No. 0;
Matches 104; Conservative 63; Mismatches 118; Indels 13; Gaps 5;

QY   1  LTNRKRNENIINALSIIPAVNSHTQMDLSPDARIEDSLCIAEGNNINPLVSASTVTGTI 60
      |||||
Db    1  MTNKRKNENIIN-----AVSNHSAQMDLLPDARIEDSLCIAEGNNIDPFVSASTVTGTI 55

QY   61 NIAGRILLGVGLVPFAGQIAISFYSLVGELWPRGRDQWEIELEHVQELINQOITENARNTA 120

```

[illegible]

Db 1135 -YTSRNRGDCAYESNSVPADYASAYEAKAYTDGRDNPCESNRGYDYLPLPAGYVTK 1193

Qy 1195 ELEYFPETDKWIEIGETEGKFIIVDSVELLME 1228

Db 1194 ELEYFPETDKWIEIGETEGKFIIVDSVELLME 1227

RESULT 14

US-08-100-709-4

Sequence 4, Application US/08100709

Patent No. 5322687

GENERAL INFORMATION:

APPLICANT: Donovan, William P.

APPLICANT: Tan, Yiping

APPLICANT: Jan, Christine S.

APPLICANT: Gonzalez Jr., Jose M.

TITLE OF INVENTION: BACILLUS THURINGIENSIS CY-ET4 AND CY-ET5

TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.

STREET: 1601 Market Street, 36th Floor

CITY: Philadelphia

STATE: Pennsylvania

COUNTRY: U.S.A.

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/100,709

FILING DATE: 19930729

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Egolf, Christopher

REGISTRATION NUMBER: 27633

REFERENCE/DOCKET NUMBER: 7205-49

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-757-1590

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1229 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-100-709-4

Query Match 80.8%; Score 5237.5; DB 1; Length 1229;

Best Local Similarity 79.9%; Pred. No. 0;

Matches 983; Conservative 94; Mismatches 149; Indels 5; Gaps 3;

Qy 1 LTSNRKNEIINALSTIPAVNSHSTQMDLSPDARIEDSLCIAEONNINPLVASTVQTGI 60

Db 1 LTSNRKNEIINALSTIPVNSPSTQMDLSPDARIEDSLCIAEONNINPLVASTVQTGI 60

Qy 61 NIAGRILGVLPAGQIASFYFLVCELWPRGDOWEIEFLEHVEQLINQOITENARNTA 120

Db 61 NIAGRILGVLPAGQIASFYFLVCELWPRGDOWEIEFLEHVEQLINQOITENARNTA 120

Qy 121 LARLOGLGDSFRAYQOQSLDLENDRDARTSRVLYTQYIALELDFLNAMPLFAIRNQVEP 180

Db 121 LARLEGLRGVRSYQQALETWLDNRDARSIIILERYVALELDITTAIPLPRIRNEVP 180

Qy 181 LLMVYAQAANLHLLLDASLFGSEFGLTSQETQRYRYRQVQOTRDSYDVCWEYATGLN 240

Db 181 LLMVYAQAANLHLLLDASLFGSEMGWASSDVNQYQEQIRYTEESYNSHCYQWYNTGLN 240

Qy 241 SLRGTNAASWVRNQFRDRLTLGLVDLVALFPSPYDTRTYPINTSAOLTREVVYTDALGATG 300

Db 241 NLRGTAESWLRNQFRDRLTLGLVDLVALFPSPYDTRTYPINTSAOLTREIYTDPIGRTN 300

Qy 301 V--NMASMMWNNNAPSFAIETAVIRSPHLLDFLEOLTIFSTSSRWSSATRHMTYWRGHT 358

Db 301 APSGFATNFWNNAPSFAIETAAIFRPHLLDFPEOLTIIYSASSRWSSATQHMYWVGH 360

Qy 359 IQSRPIGGGLNTSTHGST-NTSINPVLSPFSDRVVWTESYAGVLLMGILYLEPIHGVPY 417

Db 361 LNFPRIGGTLNTSTQGLTNTSINPVLQFTSRDVRVTESNAGTNI--LFTTPVNGVPA 418

Qy 418 RENFRPONTFERGTANYSPYSPGLQKDSSTELPPEPTERPNYSEYSHRLSHIGLIS 477

Db 419 RFNFIPQNIYERGATTYSQYQGVGQLFDSSTELPPEPTERPNYSEYSHRLSHIGLII 478

Qy 478 QSRVHVPSVSWTHRSADRTNTISSDSTIQLPVKSNLNSGTSVVSQPGFTGGDIIRTNV 537

Db 479 GNTLRAPVYSWTHRSADRTNTIGPNRITQIPVAKNLHSGVTVVGGPGFTGGDILRTN 538

Qy 538 NGSVLSGLNFNTSILQVRVRVRYAASQTMVLRTVVGSTTDFDQGPSTMSANESITSQ 597

Db 539 TGTFGDIRLINVPLSQVRVRIRYASTTDLQFTTRINGTTVNIIGNFSRTMNRGDNLEYR 598

Qy 598 SFRFAEPFVCGISASGSGTAGISISNNAGRTQTHFDKIERIPITATREAEVDLERAQAVN 657

Db 599 SFRTAGFSTPFNFNAQSTFTLGAQSFNSQEVVIDRVEFVPAEVTFEAEVDLERAQAVN 658

Qy 658 ALFTNTNPRRLKTDVTDYHIDQVSNLVACLSDFCLDEKRELEKVKYAKRLSDERNLLQ 717

Db 659 ALFTSNPRRLKTDVTDYHIDQVSNLVACLSDFCLDEKRELEKVKYAKRLSDERNLLQ 718

Qy 718 DPNFTSINKQDPFISTNBNQSFTHQSEHGWGSGENITIQEGNDVFKENYVTLPGTFN 777

Db 719 DPNFTSISQSPASIDGOSNFPISNELSEHGWGSGANVTIQEGNDVFKENYVTLPGTFN 778

Qy 778 ECVPTVLYOKIGESLKAQYTRQYLCGYIEDSODLEIYLIRYNAKHETLDVPGTESVWPLS 837

Db 779 ECVPTVLYOKIGESLKAQYTRQYLCGYIEDSODLEIYLIRYNAKHETLDVPGTESVWPLS 838

Qy 838 VESPIGRCEPNRCAPHFENWPDLCSCRDGKCAHSHHFLSDIDIGCTDLHENLGVV 897

Db 839 VESPIGRCEPNRCAPHFENWPDLCSCRDGKCAHSHHFLSDIDIGCTDLHENLGVV 898

Qy 898 VFKIKTOEGHARLGNLEFIEEKLLEALSRVGAEKWRDKREKLOLETKRYVYTAKEA 957

Db 899 VFKIKTOEGHARLGNLEFIEEKLLEALSRVGAEKWRDKREKLOLETKRYVYTAKEA 958

Qy 958 VDALFVDSQVNRLOADTNGIMHAADKLVRHIREAYLSLSVIPGVNAEIELEGRIIT 1017

Db 959 VDALFVDSQVNRLOADTNGIMHAADKLVRHIREAYLSLSVIPGVNAEIELEGRIIT 1018

Qy 1018 AISLYDARNVVKNGDFNNGLCNWKVGHVDVQOSHRSVLRVPEWEAEVSAQVRVCPGRG 1077

Db 1019 AMSLYDARNVVKNGDFNNGLCNWKVGHVDVQOSHRSVLRVPEWEAEVSAQVRVCPGRG 1078

Qy 1078 YILRVYAYKEGEGCVTHIEINNTDELKFKNCEEEVYPTDGTCTNDYTAHQGTAVCN 1137

Db 1079 YILRVYAYKEGEGCVTHIEINNTDELKFKNCEEEVYPTDGTCTNDYTAHQGTAAACN 1138

Qy 1138 SRNAGYDAYVDVTTASVNVKPYEETVTDVRDNHCEYDRGVNVYPLPAGYVTKLE 1197

Db 1139 SRNAGYDAYVDVTTASVNVKPYEETVTDVRDNHCEYDRGVNVYPLPAGYVTKLE 1198

Qy 1198 YPPTDKVWIEIGETEGKFIIVDSVELLME 1228

Db 1199 YPPTDKVWIEIGETEGKFIIVDSVELLME 1229

RESULT 15

US-08-176-865-4

Sequence 4, Application US/08176865

Patent No. 5616319

GENERAL INFORMATION:

APPLICANT: Donovan, William P.

APPLICANT: Tan, Yiping

; APPLICANT: Jany, Christine S.
 ; APPLICANT: Gonzalez Jr., Jose M.
 ; TITLE OF INVENTION: BACILLUS THURINGIENSIS clyt4 AND clyt5
 ; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Panich Schwarze Jacobs & Nadel c/o A.S.
 ; ADDRESSEE: Nadel
 ; STREET: 1601 Market Street, 36th Floor
 ; CITY: Philadelphia
 ; STATE: Pennsylvania
 ; COUNTRY: U.S.A.
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA: US/08/176,865
 ; FILING DATE: 30-DEC-1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/100,709
 ; FILING DATE: 29-JUL-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Egolf, Christopher
 ; REGISTRATION NUMBER: 27633
 ; REFERENCE/DOCKET NUMBER: 7205-49
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 215-757-1590
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1229 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-176-865-4

Query Match 80.8%; Score 5237.5; DB 1; Length 1229;
 Best Local Similarity 79.9%; Pred. No. 0;
 Matches 983; Conservative 94; Mismatches 149; Indels 5; Gaps 3;

Qy	1	LTNRKNEEIIINALSIPAVNSHSTOMDLSFDARIEDSLCIAEGNNINPLVSASTVQTGI	60
Db	1	LTNRKNEEIIINALSIPTVSNPSTOMNLSFDARIEDSLCVAEWNIDPPFVSASTVQTGI	60
Qy	61	NIAGRILGLVGPAGQIASFYSLVGLWELPRGRDQWEIIFLEHVEQLINQOITENARNTA	120
Db	61	NIAGRILGLVGPAGQIASFYSLVGLWELPGRDQWEIIFLEHVEQLIRQOQVNTENTNTA	120
Qy	121	LARLOGLGDSFRAYQQSLEDWLENRDDARTSRVLYTOYIALELDPLNAMPLFAIRNOEVP	180
Db	121	IARLEGLRGYRSYQQALETWLDNRNDARSIIILERYVALELDITTAIPLFIRNEEVP	180
Qy	181	LMVYAAQANLHLLLDASLFGSEFGLTQEIORYYERQVEQTRDYSYCVWYNTGLN	240
Db	181	LMVYAAQANLHLLLDASLFGSEWGMSSDQVYQEQIRVTEEYSNHCWQWYNTGLN	240
Qy	241	SLRGTAASVRYNQFRDRLTLGVLDLVALFPSTDTTYPINTSAQLTREYVYDAIGATG	300
Db	241	NLRGTNAESLRYNQFRDRLTLGVLDLVALFPSTDTTYPINTSAQLTREYVYDAIGATG	300
Qy	301	V--NMAAMNNYNNAPSAETAETAVIRSPLHLDLEQLTIFSTSSRWASATRHMTYRNGHT	358
Db	301	APSGFASTNNFNNAPSAETAETAVIRSPLHLDLEQLTIFSTSSRWASATRHMTYRNGHT	360
Qy	359	IQSRPIGGGLNTSTHGST--NTSINPVLSPFSDRVYVWTESVAGVLLWGIYLEPIHGVPV	417
Db	361	LNFRPIGGTNTSTOGLTNTSINPVLQFTRSDRVYVWTESNAGTNI--LFTTPVNGVPA	418
Qy	418	RFNFRNPONTFERGTANYQPYESPGQLKDSFELPPTTERPNYESYSHRLSHIGLIS	477

Db	419	RFNFINPQNIYERGATTYSQPYQGVGIQLPDSFETELPPETTERPNYESYSHRLSHIGLII	478
Qy	478	QSRVHVVPVSWTHRSADRTNTISSDSITQIPLVKSNLMSGT5VWSPGPTGDDIIRTNV	537
Db	479	GNTLRAPVTSWTHRSADRTNTIGPNRITQIPLVKALNLHSGVT5VWSPGPTGDDIIRTN	538
Qy	538	NGSVLSMGLNFNTSLQRYRVRYAASQTMVLRTVVGSTTDFDQSPSTMSANESLTSQ	597
Db	539	TGTFGDIRLNLINPVLQRYRVRYASTTDLQFFTRINGTVVIGNFSRTMNGDNLEJR	598
Qy	598	SFRFAEPFVPGISASGSOTAGISISNNAGRTQTFHDKIEFIPITATFAEAYDLERAQAVN	657
Db	599	SFRTAGSTFPFNLQAQSTFTLCAQSPSQEVVIDRVEFVPAEVTPEAYDLERAQAVN	658
Qy	658	ALFTNTNPRRLKTDVTDYHIDQVSNLVACLSDPCLDEKRELLEKVKYAKRLSDERNLLQ	717
Db	659	ALFTSTNPRRLKTDVTDYHIDQVSNLVACLSDPCLDEKRELFEKVKYAKRLSDERNLLQ	718
Qy	718	DPNFTSINKQPDFISTNEQSNFTSIHQSEHGHWGSENITIOEGNDVFKENYVTLPGTEN	777
Db	719	DPNFTFISGQSFASIDGOSNFPFISNELSEHGHWGANVTIOEGNDVFKENYVTLPGTEN	778
Qy	778	ECYPTLYQKIGSELKAYTRYQLRGYIEDSQDLIYLRIRYNAKHETLDVPGTESVWPLS	837
Db	779	ECYPNLYQKIGSELKAYTRYQLRGYIEDSQDLIYLRIRYNAKHETLDVPGTDSLWPLS	838
Qy	838	VESPIGRCEPNRCAPHFENWPDLCSDRGEKCAHSHHFSLDIDTGTDLLENLGVW	897
Db	839	VESPIGRCEPNRCAPHFENWPDLCSDRGEKCAHSHHFTLIDVGTDLLENLGVW	898
Qy	898	VFKIKTQEGHARLGNLEFTEEEKLLGENLSRVKRAEKKWDKKEKLOLETKRVVYTEAKEA	957
Db	899	VFKIKTQEGYARLGNLEFTEEEKLLGENLSRVKRAEKKWDKKEKLOLETKRVVYTEAKEA	958
Qy	958	VDALFVDSQVNRLOADTNIQMIHAADKLVRHIREAYLSLSVPGVNAEIFEELGRIIT	1017
Db	959	VDALFVDSQVQLOADTNIQMIHAADKLVRHIREAYLSLSELVPGVNAEIFEELGHIIT	1018
Qy	1018	AISLYDARNVYKNGDFNNGLACVNVKGVHDVQOSHRSVLVPIPEWEAEVSQAVRVCPRG	1077
Db	1019	AMSLYDARNVYKNGDFNNGLTCVNVKGVHDVQOSHRSVLVPIPEWEAEVSQAVRVCPRG	1078
Qy	1078	YILRVTAKEGYGEGCVTHIEIENNTDELKFKNCEEEVYPTDTGTCDNDYTAHQGTAVCN	1137
Db	1079	YILRVTAKEGYGEGCVTHIEIENNTDELKFKNCEEEVYPTDTGTCDNDYTAHQGTAAACN	1138
Qy	1138	SRNAGYEDAYEVDTTASVNYKPYEEETVTVRRDNHCEYDRGVYVNPPLPAGVYTKELE	1197
Db	1139	SRNAGYEDAYEVDTTASVNYKPYEEETVTVRRDNHCEYDRGVYVNPVPVAGVYTKELE	1198
Qy	1198	YFPETDKVWIEIGETEGKFIVDSVELLMEE	1228
Db	1199	YFPETDKVWIEIGETEGKFIVDSVELLMEE	1229

Search completed: April 21, 2005, 04:18:31
 Job time : 63 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 20:18:09 ; Search time 126 Seconds
(without alignments)
3769.379 Million cell updates/sec

Title: US-10-614-524-2
Perfect score: 6479
Sequence: 1 LTNKRNKNIINLSIPAV.....IGTEGKFIVDSVELLMEE 1228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04.*
1: Geneseqpl980s.*
2: Geneseqpl990s.*
3: Geneseqpl2000s.*
4: Geneseqpl2001s.*
5: Geneseqpl2002s.*
6: Geneseqpl2003as.*
7: Geneseqpl2003bs.*
8: Geneseqpl2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6479	100.0	1228	4 AAB84628	AAB84628 Amino aci
2	6462	99.7	1228	4 AAU02039	AAU02039 B. thurin
3	5926.5	91.5	1227	2 AAW44321	AAW44321 Bacillus
4	5926.5	91.5	1227	4 AAB19950	AAB19950 Bacillus
5	5912.5	91.3	1227	4 AAU02046	AAU02046 B. thurin
6	5740	88.6	1228	2 AAR50955	AAR50955 Bacillus
7	5436.5	83.9	1227	2 AAY31990	AAY31990 Chimeric
8	5237.5	80.8	1229	2 AAR54074	AAR54074 CryET5. 2
9	5237.5	80.8	1229	2 AAW35259	AAW35259 Bacillus
10	5237.5	80.8	1229	2 AAW17699	AAW17699 CryET5. 3
11	5237.5	80.8	1229	2 AAW87633	AAW87633 CryET5 pr
12	5237.5	80.8	1229	2 AAY30923	AAY30923 B. thurin
13	5237.5	80.8	1229	2 ADK98479	ADK98479 B. thurin
14	5233.5	80.8	1230	8 ADK98484	ADK98484 B. thurin
15	5233.5	80.8	1230	8 ADK98489	ADK98489 B. thurin
16	5233.5	80.8	1230	8 ADK98481	ADK98481 B. thurin
17	5233.5	80.8	1230	8 ADK98491	ADK98491 B. thurin
18	5233.5	80.8	1230	8 ADK98487	ADK98487 B. thurin
19	5189.5	80.1	1209	4 AAU02094	AAU02094 Bacillus
20	5183.5	80.0	1221	4 AAU00421	AAU00421 B. thurin
21	5169.5	79.8	1221	4 AAU00420	AAU00420 B. thurin
22	5108	78.8	1186	2 AAY16796	AAY16796 Amino aci
23	4791	73.9	1208	4 AAU02093	AAU02093 Bacillus
24	4303.5	66.4	1217	4 AAU02092	AAU02092 Bacillus
25	3977.5	61.4	1174	2 AAR89494	AAR89494 B.T. toxi

26	3977.5	61.4	1174	2 AAW09043	AAW09043 Bacillus
27	3974.5	61.3	1174	2 AAR25825	AAR25825 Novel tox
28	3969.5	61.3	1174	2 AAR08257	AAR08257 B.thuringi
29	3791	58.5	1176	2 AAW46737	AAW46737 Amino aci
30	3791	58.5	1176	2 AAW47035	AAW47035 Bacillus
31	3579	55.2	1170	2 AAR63231	AAR63231 Crystall p
32	3556	54.9	1167	2 AAR54073	AAR54073 CryET4. 2
33	3556	54.9	1167	2 AAW35258	AAW35258 Bacillus
34	3556	54.9	1167	2 AAW17700	AAW17700 CryET4. 3
35	3556	54.9	1167	2 AAW87632	AAW87632 CryET4. pr
36	3556	54.9	1167	2 AAY30922	AAY30922 B. thurin
37	3502.5	54.1	1189	2 AAW61345	AAW61345 a mutated
38	3502.5	54.1	1189	2 AAY17783	AAY17783 EG12111 c
39	3502.5	54.1	1189	3 AAY82431	AAY82431 Bacillus
40	3502.5	54.1	1189	5 AAE26471	AAE26471 Bacillus
41	3502.5	54.1	1189	5 ABG93849	ABG93849 B. thurin
42	3502.5	54.1	1189	7 ABW02350	ABW02350 B. thurin
43	3500.5	54.0	1189	2 AAW61335	AAW61335 A mutated
44	3500.5	54.0	1189	2 AAY17788	AAY17788 B. thurin
45	3500.5	54.0	1189	3 AAY82396	AAY82396 Bacillus

ALIGNMENTS

RESULT 1

AAB84628
ID AAB84628 standard; protein; 1228 AA.

XX AAB84628;

AC
XX 05-SEP-2001 (first entry)

DT
XX Amino acid sequence of a CryIbF insecticidal protein.

DE
XX CryIbF; insecticidal protein; CryIbJd; Cry9Fa; lepidoptera; coleoptera;
KW insect pest; transgenic plant.

XX
XX Bacillus thuringiensis.

OS
XX WO200147952-A2.

PN
XX 05-JUL-2001.

PD
XX 19-DEC-2000; 2000WO-EP013184.

PF
XX 28-DEC-1999; 99US-0173387P.

PR
XX (AVET) AVENTIS CROPS SCIENCE NV.

PA
XX Arnaut G, Boets A, Damme N, Mathieu E, Vanneste S, Van Rie J;

PI
XX WPI; 2001-425619/45.

DR
XX N-PSDB; AAB28240.

DR
XX Novel insecticidal proteins CryIbJd, Cry9Fa and CryIbF, derived from
PT Bacillus thuringiensis, useful for controlling insects in plants.

PS
XX Claim 11; Page 37-41; 65pp; English.

XX The present sequence represents CryIbF, an insecticidal protein derived
CC from Bacillus thuringiensis. The specification also describes CryIbJd and
CC Cry9Fa. The Cry proteins have activity against lepidopteran and
CC coleopteran insect pests. CryIbF, CryIbJd and Cry9Fa polynucleotides and
CC polypeptides are useful for obtaining a plant with resistance to insects.
CC Cry polynucleotides are useful for producing transgenic plants which are
CC resistant to insects

XX
SQ Sequence 1228 AA;

Query Match 100.0%; Score 6479; DB 4; Length 1228;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LTSNRKNEIINALSIPAVSNHSTOMDLPDARIEDSLCIAEGNNINPLVASTVOTGI 60
DB 1 LTSNRKNEIINALSIPAVSNHSTOMDLPDARIEDSLCIAEGNNINPLVASTVOTGI 60
QY 61 NIAGRIILGVLGVPFAGIAGFYSFLVGLMWRGRDQWEIIFLEHVEQLINQIITENARNTA 120
DB 61 NIAGRIILGVLGVPFAGIAGFYSFLVGLMWRGRDQWEIIFLEHVEQLINQIITENARNTA 120
QY 121 LARLOGLGDSFRAYQOQSLDLEWRDNDARTSRVLYTOYIALELDFLNAMPLFAIRNOEVP 180
DB 121 LARLOGLGDSFRAYQOQSLDLEWRDNDARTSRVLYTOYIALELDFLNAMPLFAIRNOEVP 180
QY 181 LLMVYAQAANLHLLLDASLFGSEFGLTQEIQRVYVEROVEDTRDYSYCVEMWNTGLN 240
DB 181 LLMVYAQAANLHLLLDASLFGSEFGLTQEIQRVYVEROVEDTRDYSYCVEMWNTGLN 240
QY 241 SLRGNTAASVVRYNQFRDLTLGLVDLVALFPSTYDTRTPINTSAQLTREVTYDAIGATG 300
DB 241 SLRGNTAASVVRYNQFRDLTLGLVDLVALFPSTYDTRTPINTSAQLTREVTYDAIGATG 300
QY 301 VNMAWMWYNNAAPSFAIETAVIRSHPLLDLFLBOLTIFSTSRWSATREMTYWRGHTIQ 360
DB 301 VNMAWMWYNNAAPSFAIETAVIRSHPLLDLFLBOLTIFSTSRWSATREMTYWRGHTIQ 360
QY 361 SRPIGGGLNTSTHGSTNTSINPRLSPFSRDVYVWTESYAGVLLMGIVLEPIHGVPTVRFN 420
DB 361 SRPIGGGLNTSTHGSTNTSINPRLSPFSRDVYVWTESYAGVLLMGIVLEPIHGVPTVRFN 420
QY 421 FRNPQNTFERGTANYQSPYSPGLQKXSETLPPETTERPNYESYSHRLSHIGLISQSR 480
DB 421 FRNPQNTFERGTANYQSPYSPGLQKXSETLPPETTERPNYESYSHRLSHIGLISQSR 480
QY 481 VHPVYVSWTHRSADRTNTISDSITQIPLVKSFNLSGTSVWSPGPGTGGDIIRTNVNGS 540
DB 481 VHPVYVSWTHRSADRTNTISDSITQIPLVKSFNLSGTSVWSPGPGTGGDIIRTNVNGS 540
QY 541 VLSMGLNFNTSLQRYVRVRYAASQTMVLRTVVGSTTFDQGPSTMSANESLTSQSR 600
DB 541 VLSMGLNFNTSLQRYVRVRYAASQTMVLRTVVGSTTFDQGPSTMSANESLTSQSR 600
QY 601 FAFPPVGISASGSGTAGISISNAGRTQFHFDFKIEFIPITATFEAEYDLERAQEVNALLF 660
DB 601 FAFPPVGISASGSGTAGISISNAGRTQFHFDFKIEFIPITATFEAEYDLERAQEVNALLF 660
QY 661 TTNTPRLKTDVTDYHIDQVSNLVACLSDFCLDEKRELLKVKYAKRLSDERNLLODPN 720
DB 661 TTNTPRLKTDVTDYHIDQVSNLVACLSDFCLDEKRELLKVKYAKRLSDERNLLODPN 720
QY 721 FTSINKQPDFISTNEQSNFTSIHQSEHGHWGSGENITIQEGNDVFKENYVTLPTFNECY 780
DB 721 FTSINKQPDFISTNEQSNFTSIHQSEHGHWGSGENITIQEGNDVFKENYVTLPTFNECY 780
QY 781 PTLYQXIGISELKAITYRQYLRGYIEDSQDLIELYIRYNAKHETLDVPGTESVMPLSVES 840
DB 781 PTLYQXIGISELKAITYRQYLRGYIEDSQDLIELYIRYNAKHETLDVPGTESVMPLSVES 840
QY 841 PIGRCGPNCAPHFENWPDLCSDRGKCAHSHHFSLDIDIGCTDLHENLGVWVVK 900
DB 841 PIGRCGPNCAPHFENWPDLCSDRGKCAHSHHFSLDIDIGCTDLHENLGVWVVK 900
QY 901 IKTOEGHARLGNLEFIEEKLPGALSRVKRAEKKWDRKLEKQLETKRYVYVTEAKEAVDA 960
DB 901 IKTOEGHARLGNLEFIEEKLPGALSRVKRAEKKWDRKLEKQLETKRYVYVTEAKEAVDA 960
QY 961 LFDVDSQYRNLQADNTNIGMHAADKLVHRIEAYLSVLPVGNABEIELEGRIITAIS 1020
DB 961 LFDVDSQYRNLQADNTNIGMHAADKLVHRIEAYLSVLPVGNABEIELEGRIITAIS 1020
QY 1021 LYDARNVVKNGDFNGLACNVKGVHDVQOQSHRSVLVIPWEAEVSVQVRVCPGRGIYL 1080
DB 1021 LYDARNVVKNGDFNGLACNVKGVHDVQOQSHRSVLVIPWEAEVSVQVRVCPGRGIYL 1080

QY 1081 RVTAAYKEGYGEGCVTIHEIENNTDELKFKNCEBEEVYPTDTGTCNDYTAHQGTAVCNSRN 1140
DB 1081 RVTAAYKEGYGEGCVTIHEIENNTDELKFKNCEBEEVYPTDTGTCNDYTAHQGTAVCNSRN 1140
QY 1141 AGYEDAYEVDVTTASVNYKPTYEBEETVTVRRDNHCEYDRGYVNYPPPLPAGYMTKELEYFP 1200
DB 1141 AGYEDAYEVDVTTASVNYKPTYEBEETVTVRRDNHCEYDRGYVNYPPPLPAGYMTKELEYFP 1200
QY 1201 ETDKVMIEIGETGKPFIVDSVELLMEE 1228
DB 1201 ETDKVMIEIGETGKPFIVDSVELLMEE 1228
RESULT 2
AAU02039 standard; protein; 1228 AA.
XX AAU02039;
XX AC AC
XX DT DT
XX 29-AUG-2001 (first entry)
DE B. thuringiensis toxic crystal protein, CryET64.
XX Delta endotoxin; Lepidopteran-active; crystal protein; insecticide;
KW transgenic plant; corn; wheat; soybean; oat; cotton; rice; yre; sorghum;
KW sugarcane; tomato; tobacco; kapok; flax; potato; barley; turf grass;
KW pasture grass; berry; fruit; legume; vegetable; ornamental plant; shrub;
KW cactus; tree call; gypsy moth; looper; tobacco budworm; spruce budworm;
KW cotton leaf perforator; CryET64.
XX Bacillus thuringiensis.
OS
XX WO200119859-A2.
XX PD 22-MAR-2001.
XX PF 13-SEP-2000; 2000WO-US025361.
XX PR 15-SEP-1999; 99US-0153995P.
XX PA (MONS) MONSANTO CO.
XX Baum JA, Chu C, Donovan WP, Gilmer AJ, Rupar MJ;
XX WPI; 2001-281518/29.
XX N-PSDB; AAS02482.
PT Lepidopteran-active Bacillus thuringiensis delta-endotoxin polypeptides
PT and the polynucleotides that encode them, useful for increasing the
PT insect resistance of plant.
XX Claim 5; Page 139-141; 173pp; English.
XX The sequence represents a B. thuringiensis Lepidopteran-active delta-
CC endotoxin, crystal protein CryET64. The Lepidopteran-active B.
CC thuringiensis delta-endotoxin polypeptides may be used as compositions
CC that are applied to plant crops to protect them from insect damage. The
CC polynucleotides may be used in the production of transgenic plants that
CC express the insecticidal polypeptides and consequently have improved
CC insect resistance compared to non-transformed plants. Monocotyledonous or
CC dicotyledonous plants may be protected in this way, for example corn,
CC wheat, soybean, oat, cotton, rice, yre, sorghum, sugarcane, tomato,
CC tobacco, kapok, flax, potato, barley, turf grass, pasture grass, berry,
CC fruit, legume, vegetable, ornamental plant, shrub, cactus and/or tree
CC cell. A wide range of insects (e.g. gypsy moth, looper, tobacco budworm,
CC cotton leaf perforator and spruce budworm) may be affected by application
CC of the insecticidal polypeptides (full details given in specification)
XX
SQ Sequence 1228 AA;
Query Match 99.7%; Score 6462; DB 4; Length 1228;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1223; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LTSNRKNEIINALSTPAVNSHSTOMDLSPDARIEDSLCIAEGNNINPLVSASTVGTGI 60
 Db :|||||
 QY 1 MTSNRKNEIINALSTPAVNSHSTOMDLSPDARIEDSLCIAEGNNINPLVSASTVGTGI 60
 Db :|||||
 QY 61 NIAGRILGVLPAGPAGIAGFYFLVCELWPRGRDQWEI FLEHVEQLINQOITENARNTA 120
 Db :|||||
 QY 61 NIAGRILGVLPAGPAGIAGFYFLVCELWPRGRDQWEI FLEHVEQLINQOITENARNTA 120
 Db :|||||
 QY 121 LARLQGLGDSFRAYQQSLEDWLENRRDARTSRVLYTOYIALALEDFLNAMEPLFAIRNOEVP 180
 Db :|||||
 QY 121 LARLQGLGDSFRAYQQSLEDWLENRRDARTSRVLYTOYIALALEDFLNAMEPLFAIRNOEVP 180
 Db :|||||
 QY 181 LMVYAQAANLHLLLDASLFGSEFGLTSQEIORYYERQVEQTRDYSYCVIEWYNTGLN 240
 Db :|||||
 QY 181 LMVYAQAANLHLLLDASLFGSEFGLTSQEIORYYERQVEQTRDYSYCVIEWYNTGLN 240
 Db :|||||
 QY 241 SLRGTAASWVRYNQFRDRTGLVLDVALFPSPYDTRTYPINTSAQLTRREYVYDAIGATG 300
 Db :|||||
 QY 241 SLRGTAASWVRYNQFRDRTGLVLDVALFPSPYDTRTYPINTSAQLTRREYVYDAIGATG 300
 Db :|||||
 QY 301 VNMAWMYNNNAPSFAIAETAVIRSPHLLDFLEQLTIFSTSRWSATRHMTYWRGHTIQ 360
 Db :|||||
 QY 301 VNMAWMYNNNAPSFAIAETAVIRSPHLLDFLEQLTIFSTSRWSATRHMTYWRGHTIQ 360
 Db :|||||
 QY 361 SRPIGGGLNTSTHGSTNTSINPRLSPFSDRVYWTESYAGVLLWGIYLEPIHGVPTVRFN 420
 Db :|||||
 QY 361 SRPIGGGLNTSTHGSTNTSINPRLSPFSDRVYWTESYAGVLLWGIYLEPIHGVPTVRFN 420
 Db :|||||
 QY 421 FRNPQNTFERGTANYSPYSPGLQKDSLTPETTERPNYESYSHRLSHGLISQSR 480
 Db :|||||
 QY 421 FRNPQNTFERGTANYSPYSPGLQKDSLTPETTERPNYESYSHRLSHGLISQSR 480
 Db :|||||
 QY 481 VHPVVSWTHRSADRTNTIISDSTIQPLVKSFNLNSGTSVSGPGFTGGDIIRTNVNGS 540
 Db :|||||
 QY 481 VHPVVSWTHRSADRTNTIISDSTIQPLVKSFNLNSGTSVSGPGFTGGDIIRTNVNGS 540
 Db :|||||
 QY 541 VLSMGLNFNNTSLQRYRVRVRYAASQTMVLRVTVGSGTTDFDQGPPTMSANESLTSQSPR 600
 Db :|||||
 QY 541 VLSMGLNFNNTSLQRYRVRVRYAASQTMVLRVTVGSGTTDFDQGPPTMSANESLTSQSPR 600
 Db :|||||
 QY 601 FAEPFVGISASGQTAGISISNNAGROTFFDKIEPIITATPEAYDLERAQEAVALNF 660
 Db :|||||
 QY 601 FAEPFVGISASGQTAGISISNNAGROTFFDKIEPIITATPEAYDLERAQEAVALNF 660
 Db :|||||
 QY 661 TTNTPRELKTDVTDYHIDQVSNLVACLSDEFCLDEKRELLEKVKYAKRLSDERNLLQDPN 720
 Db :|||||
 QY 661 TTNTPRELKTDVTDYHIDQVSNLVACLSDEFCLDEKRELLEKVKYAKRLSDERNLLQDPN 720
 Db :|||||
 QY 721 FTSINKQPDFISTNEQSNFTSIHQSEHGMWGSNITIQEGNDVFKENYVTLPGTFNECY 780
 Db :|||||
 QY 721 FTSINKQPDFISTNEQSNFTSIHQSEHGMWGSNITIQEGNDVFKENYVTLPGTFNECY 780
 Db :|||||
 QY 781 PTYLYQKIGSELKAYTRYQLRGHIEDSQDLEIYLIRYNAKHETLDVPGETSVWPLSVES 840
 Db :|||||
 QY 781 PTYLYQKIGSELKAYTRYQLRGHIEDSQDLEIYLIRYNAKHETLDVPGETSVWPLSVES 840
 Db :|||||
 QY 841 PIGCGEPNRCAPHFENWPDLCSCRDGKCAHSHHFSLDIDIGCTDLHENLGVWVFX 900
 Db :|||||
 QY 841 PIGCGEPNRCAPHFENWPDLCSCRDGKCAHSHHFSLDIDIGCTDLHENLGVWVFX 900
 Db :|||||
 QY 901 IKTOEGHARLGNLEFIEBKPLLGALSRVKAEEKKRDREKLEKRYVYVTEAKEAVDA 960
 Db :|||||
 QY 901 IKTOEGHARLGNLEFIEBKPLLGALSRVKAEEKKRDREKLEKRYVYVTEAKEAVDA 960
 Db :|||||
 QY 961 LFVDSQYNRLQADTNIGMIHAADKLVHRIREAYLSLSVPGVNAEIPFEELEGRITTAIS 1020
 Db :|||||
 QY 961 LFVDSQYNRLQADTNIGMIHAADKLVHRIREAYLSLSVPGVNAEIPFEELEGRITTAIS 1020
 Db :|||||
 QY 1021 LYDARNVVKNGDFNGLACNVKGVHDVQOSSHRSVLVPEWEAEVQAVRVCPGRCYIL 1080
 Db :|||||
 QY 1021 LYDARNVVKNGDFNGLACNVKGVHDVQOSSHRSVLVPEWEAEVQAVRVCPGRCYIL 1080
 Db :|||||

QY 1081 RVTAYKEGYGEGCVTTHEIENNTDELKFKNCEBEEVYPTDGTGTCNDYTAHQGTAVCNRSN 1140
 Db :|||||
 QY 1081 RVTAYKEGYGEGCVTTHEIENNTDELKFKNCEBEEVYPTDGTGTCNDYTAHQGTAVCNRSN 1140
 Db :|||||
 QY 1141 AGYEDAYEVDTTASVNVKPYEETVTDVRRDNHCEYDRGVVNYPPPLPAGYMTKELEYFP 1200
 Db :|||||
 QY 1141 AGYEDAYEVDTTASVNVKPYEETVTDVRRDNHCEYDRGVVNYPPPLPAGYMTKELEYFP 1200
 Db :|||||
 QY 1201 ETDKWIEIGETEGKTFIVDSVELLMEE 1228
 Db :|||||
 QY 1201 ETDKWIEIGETEGKTFIVDSVELLMEE 1228
 Db :|||||
 RESULT 3
 AAW44321
 ID AAW44321 standard; protein; 1227 AA.
 XX
 AC AAW44321;
 XX
 DT 27-MAY-1998 (first entry)
 XX
 DE Bacillus thuringiensis 158C2c toxin.
 XX
 KW Bacillus thuringiensis; 158C2c; toxin; lepidopteran-active; insect;
 KW microbe; transgenic plant; resistant.
 XX
 OS Bacillus thuringiensis.
 XX
 PN US5723758-A.
 XX
 PD 03-MAR-1998.
 XX
 PF 23-MAY-1995; 95US-00448170.
 XX
 PR 13-SEP-1991; 91US-00759247.
 PR 01-JUN-1993; 93US-00069902.
 XX
 PA (MYCO) MYCOGEN CORP.
 XX
 PI Stelman S, Narva KE, Cummings DA, Payne J, Cannon RJ;
 DR WPI: 1998-178600/16.
 DR N-PSDB; AAV15222.
 XX
 PT DNA encoding Bacillus thuringiensis toxin proteins - for producing
 transgenic plants resistant to attack by lepidopteran pests.
 XX
 PS Claim 1; Col 33-40; 25pp; English.
 XX
 CC The present sequence represents a 158C2c toxin from Bacillus
 thuringiensis. The present invention describes isolated nucleic acid
 sequences encoding toxins active against lepidopteran insects. The
 present invention also describes a recombinant host transformed with a
 nucleic acid sequence encoding a toxin. The nucleic acid sequence can be
 amplified. The product is used for producing transgenic plants resistant
 to attack by lepidopteran pests. N.B. The numbers in the sequence listing
 do not correspond to the SEQ ID NO.'s mentioned in the specification e.g.
 in the specification SEQ ID NO:3, 4, 5 and 6 are said to correspond to
 the nucleic acid sequences encoding the toxins, and SEQ ID NO:7, 8, 9 and
 10 are said to correspond to the amino acid sequences of the toxins,
 whereas in the sequence listing the amino acid sequence is given after
 the nucleic acid sequence encoding it (i.e. SEQ ID NO:3 encodes SEQ ID
 NO:4)
 XX
 SQ Sequence 1227 AA;
 : Query Match 91.5%; Score 5926.5; DB 2; Length 1227;
 Best Local Similarity 91.8%; Pred. No. 0;
 Matches 1129; Conservative 36; Mismatches 60; Indels 5; Gaps 3;
 QY 1 LTSNRKNEIINALSTPAVNSHSTOMDLSPDARIEDSLCIAEGNNINPLVSASTVGTGI 60
 Db 1 LTSNRKNEIINALSTPAVNSHSTOMDLSPDARIEDSLCIAEGNNIDPFVSASTVGTGI 60

QY 61 NIAGRIILGVLPVAGPAGIAGFYSFLVGLMWRGRDOWEIFLEHVEQLINOQITENARNTA 120
DB 61 NIAGRIILGVLPVAGPAGIAGFYSFLVGLMWRGRDOWEIFLEHVEQLINOQITENARNTA 120
QY 121 LARLOGLGDSFRAYQOQSLDLEWLNDRDARTSRVLYTQYIALEDFLNAMEPLFAIRNOQV 180
DB 121 LARLOGLGDSFRAYQOQSLDLEWLNDRDARTSRVLYTQYIALEDFLNAMEPLFAIRNOQV 180
QY 181 LLMYAQAANLHLLLDASLFGSEFGITSOEIQRYERQVETQTRDYSDCVIEWYNTGLN 240
DB 181 LLMYAQAANLHLLLDASLFGSEFGITSOEIQRYERQVETQTRDYSDCVIEWYNTGLN 240
QY 241 SLRGTAASVRYNQFRDLTLGLVDLVALFPSPYDTRTYPINTSAQLTRREYVYDDAIGATG 300
DB 241 NLRGTNAESWLRYNQFRDLTLGLVDLVALFPSPYDTRTYPINTSAQLTRREYVYDDAIGATG 300
QY 301 V--NMASMNWNNAPSFSAIETAVIRSPHLLDFLEQLTIFSTSSRWSTRAHMYWRIGHT 358
DB 301 APSGFSTNFWNNAPSFSAIETAVIRSPHLLDFLEQLTIFSTSSRWSTRAHMYWRIGHT 358
QY 359 IQSPICGGLNTSHGNTSINPRLSPFRSDVYWTESVAGULLWGIYLEPHGVPTVR 418
DB 359 IQSPICGGLNTSHGNTSINPRLSPFRSDVYWTESVAGULLWGIYLEPHGVPTVR 418
QY 361 LESRTIRGSLSTHGNTSINPRLSPFRSDVYWTESVAGULLWGIYLEPHGVPTVR 418
DB 361 LESRTIRGSLSTHGNTSINPRLSPFRSDVYWTESVAGULLWGIYLEPHGVPTVR 418
QY 419 FNFRNPONTFERGTANYSPQYSPGLQKDSSETLPPETTERPNYESYSHRLSHIGLISQ 478
DB 419 FNFRNPONTFERGTANYSPQYSPGLQKDSSETLPPETTERPNYESYSHRLSHIGLISQ 478
QY 479 SRHVVPYVSWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVVGPGFTGGDIIRTNVN 538
DB 479 SRHVVPYVSWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVVGPGFTGGDIIRTNVN 538
QY 539 GSVLSMGLNFNTSLQRYVRVRYAASQTMVLRVTVGSGTTFDQGPSTMSANESLTSQS 598
DB 539 GSVLSMGLNFNTSLQRYVRVRYAASQTMVLRVTVGSGTTFDQGPSTMSANESLTSQS 598
QY 599 FRFAEFPVGISASGQTAGISINNAQRTQFHDKIEFIPITATFAEYDLERAQEAUNA 658
DB 599 FRFAEFPVGISASGQTAGISINNAQRTQFHDKIEFIPITATFAEYDLERAQEAUNA 658
QY 659 LFTNTNPRRLKTDVTDHIDOVNVLVACLSDBFCLDEKRELLKVKYAKRLSDERNLLQD 718
DB 659 LFTNTNPRRLKTDVTDHIDOVNVLVACLSDBFCLDEKRELLKVKYAKRLSDERNLLQD 718
QY 719 PNFTSINKQPDFISTNEQSNFTSHEQSEHGWSGSENIITIOEGNDVPEKENYVTLPGTFNE 778
DB 719 PNFTSINKQPDFISTNEQSNFTSHEQSEHGWSGSENIITIOEGNDVPEKENYVTLPGTFNE 778
QY 779 CYPTLYQKIGESLKAATRYQLRGYTEDSQDLLEIYLIRNAKHETLDVPGTESVWPLSV 838
DB 779 CYPTLYQKIGESLKAATRYQLRGYTEDSQDLLEIYLIRNAKHETLDVPGTESVWPLSV 838
QY 839 ESPIGRCGEPRNRCAPHEWNPDLDCSCRDGCKCAHSHHPSLDDIGCTDLHENLGVVW 898
DB 839 ESPIGRCGEPRNRCAPHEWNPDLDCSCRDGCKCAHSHHPSLDDIGCTDLHENLGVVW 898
QY 899 FKIKTQEGHARLGNLEFIEBKPLGELALSRVKAEEKWRDKREKLOLETGRVYTEAKEAV 958
DB 899 FKIKTQEGHARLGNLEFIEBKPLGELALSRVKAEEKWRDKREKLOLETGRVYTEAKEAV 958
QY 959 DALFVDSQYKRLQADTNIGMHAADKLVRHIREAYLSLSVIFGVNAEIPFEELEGRITTA 1018
DB 959 DALFVDSQYKRLQADTNIGMHAADKLVRHIREAYLSLSVIFGVNAEIPFEELEGRITTA 1018
QY 1019 ISLYDARNVKGDFNNGLAQWNVKGVHDVQOQSHRSVLVPEWEAEVSAQVRCVPGRGY 1078
DB 1019 ISLYDARNVKGDFNNGLAQWNVKGVHDVQOQSHRSVLVPEWEAEVSAQVRCVPGRGY 1078
QY 1079 ILRVTAKEGVGCGCVTHIEIENNTDRLKPKNCEEEVEVPTDGTGNDYTAHOGTAVCNS 1138
DB 1079 ILRVTAKEGVGCGCVTHIEIENNTDRLKPKNCEEEVEVPTDGTGNDYTAHOGTAVCNS 1138

QY 1139 RNAGYEDAYEVDTTASVNYKPTVEEETTYDVRDNRHCEYDRGVNYPPLPAGYMTKELEY 1198
DB 1139 RNAGYEDAYEVDTTASVNYKPTVEEETTYDVRDNRHCEYDRGVNYPPLPAGYMTKELEY 1198
QY 1199 FPETDKVWIBIGETEGKFIVDSVELLMEE 1228
DB 1199 FPETDKVWIBIGETEGKFIVDSVELLMEE 1228
RESULT 4
AAB19950
ID AAB19950 standard; protein; 1227 AA.
XX AAB19950;
AC AAB19950;
XT 19-MAR-2001 (first entry)
DT 19-MAR-2001 (first entry)
XX
DE Bacillus thuringiensis delta-endotoxin 1582C2.
XX Delta-endotoxin; 1582C2; toxin; lepidoptera; crystal protein;
KW transgenic plant; insect resistance; crop protection; biological control.
XX
OS Bacillus thuringiensis.
XX US6150589-A.
PN 21-NOV-2000.
PD 31-OCT-1997; 97US-00961803.
XX 23-MAY-1995; 95US-00448170.
PR (MYCO) MYCOGEN CORP.
PA Narva KE, Stelman S, Payne J, Cummings DA, Cannon RJC;
PI WPI; 2001-049107/06.
XX N-PSDB; AAA89221.
DR New genes isolated from Bacillus thuringiensis which encode toxins
XX against lepidopteran insects and which can be used to transform various
PT hosts to express the B.t. toxin and confer insect resistance.
XX Example 2; Col 35-42; 25pp; English.
XX The present sequence is that of the C-terminal region of novel delta-
CC endotoxin 1582C2 of Bacillus thuringiensis (B.t.) isolate PS158C2 (NRRL B
CC -18872). The amino acid sequence is predicted from a partial gene
CC sequence for 1582C2 (see AAA89222) obtained from a PS158C2 gene library.
CC The invention provides novel B.t. delta-endotoxins 1582Ca-d (see AAB19947
CC -50), which are active against lepidopteran pests. Expression of the
CC delta-endotoxins in transformed plants confers resistance to these insect
CC pests, while transformed microorganisms can be applied to sites where
XX they will proliferate and be ingested by the pest
SQ Sequence 1227 AA;

Query Match 91.5%; Score 5926.5; DB 4; Length 1227;
Best Local Similarity 91.8%; Pred. No. 0;
Matches 1129; Conservative 36; Mismatches 60; Indels 5; Gaps 3;
QY 1 LTSNRKNEININALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
DB 1 LTSNRKNEININALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
QY 61 NIAGRILGVLPVAGIAGFYSFLVGLMWRGRDOWEIFLEHVEQLINOQITENARNTA 120
DB 61 NIAGRILGVLPVAGIAGFYSFLVGLMWRGRDOWEIFLEHVEQLINOQITENARNTA 120
QY 121 LARLOGLGDSFRAYQOQSLDLEWLNDRDARTSRVLYTQYIALEDFLNAMEPLFAIRNOQV 180
DB 121 LARLOGLGDSFRAYQOQSLDLEWLNDRDARTSRVLYTQYIALEDFLNAMEPLFAIRNOQV 180

```
QY 181 LLVYAAANLHLLLDASLFGSEGLTQSOIQRVYERQVQTRDYSKYCVENYNTGIN 240
Db 181 LLVYAAANLHLLLDASLFGSEGLTQSOIQRVYERQVQTRDYSKYCVENYNTGLN 240
QY 241 SLRGTNAASVRYNQFRDRLTLGLVDLVALFPSPDYTRTYPINTSAQLTRVYTDGATG 300
Db 241 NLRGTAESWLRYNQFRDRLTLGLVDLVALFPSPDYTRTYPINTSAQLTRVYTDPIGRTN 300
QY 301 V--NMASMNWNNAPSAISAETAVIRSPHLLDFLEQLTTFSTSSRWSATRHTMYWRIGHT 358
Db 301 APSGFASSTNFWNNAPSAISAETAAVIRPPHLLDFPEQLTTFVSLRSWSNTQYNNYWGHR 360
QY 359 IQSRPIGGGLNTSHGTSNTSINPVLRSFSDRVYWTESYAGVLLGWILEPIHGVTVR 418
Db 361 LERKTIGSLTSHGTSNTSINPVLRSFSDRVYWTESYAGVLLGWILEPIHGVTVR 418
QY 419 FNRPNQPTFRGTANYSQPYSPGLQKDSFELPPTETTERPNYESYSHRLSHIGLISQ 478
Db 419 FNRPNPLSL-RGSLLYTIGYTGVTQLFDSFETELPPTETTERPNYESYSHRLSNIRLISG 477
QY 479 SRVHPVYVSWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVVGPGFTGGDIIRTNVN 538
Db 478 NTLRAPVYVSWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVVGPGFTGGDIIRTNVN 537
QY 539 GSVLSMGLNFNTSLQRYRVRYAASQTMVLRVTVCGSTTFDQGPSTMSANESLTSQS 598
Db 538 GSVLSMGLNFNTSLQRYRVRYAASQTMVLRVTVCGSTTFDQGPSTMSANESLTSQS 597
QY 599 FRAEPFVPGISASGSQTAGISINNAQRQTFHFDKIEFIPITATFEAEYDLERAQAVNA 658
Db 598 FRAEPFVPGISASGSQTAGISINNAQRQTFHFDKIEFIPITATFEAEYDLERAQAVNA 657
QY 659 LFTNTPRRLKTDVTDVHIDQVSNLACLDSFCLDEKRELLKVKYAKRLSDERNLLQD 718
Db 658 LFTNTPRRLKTDVTDVHIDQVSNLACLDSFCLDEKRELLKVKYAKRLSDERNLLQD 717
QY 719 PNFTSINKQPDFSTNEQSFNTSHEQSEHGWSGSENIITQEGNDVPKENYVILPGTFNE 778
Db 718 PNFTSINKQPDFSTNEQSFNTSHEQSEHGWSGSENIITQEGNDVPKENYVILPGTFNE 777
QY 779 CYPTLYQKIGESLKAATYQLRGYIEDSQDLEIYLIRNAXHETLDVPGTSVWPLSV 838
Db 778 CYPTLYQKIGESLKAATYQLRGYIEDSQDLEIYLIRNAXHETLDVPGTSVWPLSV 837
QY 839 ESPIGRCGEPRNRCAPHEWNPDLDCSDRGEKCAHSHHFSLDIDIGCTDLHENLGVVV 898
Db 838 ESPIGRCGEPRNRCAPHEWNPDLDCSDRGEKCAHSHHFSLDIDIGCTDLHENLGVVV 897
QY 899 FKIQTQEGHARLGNLEFIEBKPLIGALSRVKAERKWRDREKLQLETGRVYTEAKEAV 958
Db 898 FKIQTQEGHARLGNLEFIEBKPLIGALSRVKAERKWRDREKLQLETGRVYTEAKEAV 957
QY 959 DALFVDSQYNRLQADNTNIGMHAADKLVRIRIAYLSLSVPIGVNAIEFEELEGRITA 1018
Db 958 DALFVDSQYNRLQADNTNIGMHAADKLVRIRIAYLSLSVPIGVNAIEFEELEGRITA 1017
QY 1019 ISLYDARNVYKNGDFNNGLACWNKGVHDVQSHHRSVLVPIPEWAEVSQAVRVCGRGY 1078
Db 1018 ISLYDARNVYKNGDFNNGLACWNKGVHDVQSHHRSVLVPIPEWAEVSQAVRVCGRGY 1077
QY 1079 ILRVYAYKEGEGCVTIHIEIENNTDLKFNCEEEVYPTDGTGNDYTAHQGTAVCNIS 1138
Db 1078 ILRVYAYKEGEGCVTIHIEIENNTDLKFNCEEEVYPTDGTGNDYTAHQGTAVCNIS 1137
QY 1139 RNAGYDAYEVDVTASVNYKPTVEEETDVRDNDHCEYDRGVVNPPLPAGVMTKELEY 1198
Db 1138 RNAGYDAYEVDVTASVNYKPTVEEETDVRDNDHCEYDRGVVNPPLPAGVMTKELEY 1197
QY 1199 FPETDKWVIEIGETEGKFIIVDSVELLLMEE 1228
Db 1198 FPETDKWVIEIGETEGKFIIVDSVELLLMEE 1227
```

```
RESULT 5
AAU02046
ID AAU02046 standard; protein; 1227 AA.
XX
AC AAU02046;
XX
DT 29-AUG-2001 (first entry)
XX
DE B. thuringiensis toxic crystal protein, CryET54.
XX
KW Delta endotoxin; Lepidopteran-active; crystal protein; insecticide;
KW transgenic plant; corn; wheat; soybean; oat; cotton; rice; sorghum;
KW sugarcane; tomato; tobacco; kapok; flax; potato; barley; turf grass;
KW pasture grass; berry; fruit; legume; vegetable; ornamental plant; shrub;
KW cactus; tree cell; gypsy moth; looper; tobacco budworm; spruce budworm;
KW cotton leaf perforator; CryET54.
XX
OS Bacillus thuringiensis.
XX
PN WO200119859-A2.
XX
PD 22-MAR-2001.
XX
PF 13-SEP-2000; 2000WO-US025361.
XX
PR 15-SEP-1999; 99US-0153995P.
XX
PA (MONS ) MONSANTO CO.
XX
PI Baum JA, Chu C, Donovan WP, Gilmer AJ, Rupar MJ;
XX
DR WPI; 2001-281518/29.
XX
DR N-PSDB; AAS02489.
XX
PT Lepidopteran-active Bacillus thuringiensis delta-endotoxin polypeptides
PT and the polynucleotides that encode them, useful for increasing the
PT insect resistance of plant.
XX
PS Claim 17; Page 169-173; 173pp; English.
XX
CC The sequence represents B. thuringiensis Lepidopteran-active delta-
CC endotoxin, crystal protein CryET54. The Lepidopteran-active B.
CC thuringiensis delta-endotoxin polypeptides may be used as compositions
CC that are applied to plant crops to protect them from insect damage. The
CC polynucleotides may be used in the production of transgenic plants that
CC express the insecticidal polypeptides and consequently have improved
CC insect resistance compared to non-transformed plants. Monocotyledonous or
CC dicotyledonous plants may be protected in this way, for example corn,
CC wheat, soybean, oat, cotton, rice, rye, sorghum, sugarcane, tomato,
CC tobacco, kapok, flax, potato, barley, turf grass, pasture grass, berry,
CC fruit, legume, vegetable, ornamental plant, shrub, cactus and/or tree
CC cell. A wide range of insects (e.g. gypsy moth, looper, tobacco budworm,
CC cotton leaf perforator and spruce budworm) may be affected by application
CC of the insecticidal polypeptides (full details given in specification).
CC Note: The present sequence does not have a cryET number assigned
CC explicitly in the specification, cryET54 (see table 5, page 56) is not
CC assigned to any Seq.ID number, the indexer has matched the spare sequence
CC to the spare cryET number, which may be incorrect
XX
SQ Sequence 1227 AA;
```

```
Query Match 91.3%; Score 5912.5; DB 4; Length 1227;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1127; Conservative 36; Mismatches 62; Indels 5; Gaps 3;
QY 1 LTSNRKNEIINALSIPAVSNHSTOMDLSPDARIEDSLCIAEGNNINPLVASTVQTGI 60
Db 1 LTSNRKNEIINALSIPAVSNHSAQNLSLSTDIEDSLCIAEGNNIDPFPVASTVQTGI 60
QY 61 NIAGRILGVLPVFPAGQIASFYSLVGCFLWPRGRDQWEIFLEHVQOLINQITENARNTA 120
Db 61 NIAGRILGVLPVFPAGQIASFYSLVGCFLWPRGRDQWEIFLEHVQOLINQITENARNTA 120
```


QY 121 LARLOGLGDSFRAYQOQSLDLEWLNDRDARTSRVLYTQYIALELDFLNAMPFLAIRNOEVP 180
 Db 121 LARLOGLGNSFRAYQOQSLDLEWLNDRDARTSRVLYTQYIALELDFLNAMPFLAIRNOEVP 180
 QY 181 LLMVYQAANLHLLLDASLFGSEFGITQEIORYYERQVEQTRDYSYCVWEYNTGLN 240
 Db 181 LLMVYQAANLHLLLDASLFGSEFGITQEIORYYERQVEQTRDYSYCVWEYNTGLN 240
 QY 241 SLRGTNAASWVRVNOFRDLTLGVLDLVALFPSPYDTRTYPINTSAQLTRVYTDALCATG 300
 Db 241 NLRGTAESWLRVNOFRDLTLGVLDLVALFPSPYDTRTYPINTSAQLTRVYTDALCATG 300
 QY 301 V--NMASNNWNNAPSAIETAVIRSPHLLDFLEQLTIFTSRSSWATRHMTYWRGHT 358
 Db 301 APSGFASNNWNNAPSAIETAVIRSPHLLDFLEQLTIFTSRSSWATRHMTYWRGHT 358
 QY 359 IQSRPIGGGLTSHGNTSINPVRSLPFRSDVYWTESYAGVLLWGLIYEPHIGVPTVR 418
 Db 361 LESTIRGSLTWHGNTSINPVTLOFTSRDYRTESFAGINI--LLTPVNGVFWAR 418
 QY 419 FNRPNPONTFRGTANYSPQVSPGLQKDSFELPPTERRPNYSYSHRLSHIGLISQ 478
 Db 419 FNRPNPLNSL--RGLLTIGYTGVTQTFSETELPPTERRPNYSYSHRLSNIRLISG 477
 QY 479 SRVHPVYVWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVVGSGFTGGDIIRTNVN 538
 Db 478 NTLRAPVYVWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVVGSGFTGGDIIRTNVN 537
 QY 539 GSVLSMGLNFNNTSLQRYRVRYAASQTVLWLVTVGGSTTFDQGPSTMSANESLTSQS 598
 Db 538 GSVLSMGLNFNNTSLQRYRVRYAASQTVLWLVTVGGSTTFDQGPSTMSANESLTSQS 597
 QY 599 FRAEPFVGISASGSGTAGISINNAQRTFFPKIEFIPITATFAEYDLERAQAVNA 658
 Db 598 FRAEPFVGISASGSGTAGISINNAQRTFFPKIEFIPITATFAEYDLERAQAVNA 657
 QY 659 LFTNTPRRLKTDVTHIDQVNLVACLSDFCLEKRELLKVKYAKLSDERNLLQD 718
 Db 658 LFTNTPRRLKTDVTHIDQVNLVACLSDFCLEKRELLKVKYAKLSDERNLLQD 717
 QY 719 PNFTSINKQPDFSTNEOSNFTSIHSEHGWSGSENITIOEGNDVFKENYVTLPGTFNE 778
 Db 718 PNFTSINKQPDFSNNEQSNFTSIHSEHGWSGSENITIOEGNDVFKENYVTLPGTFNE 777
 QY 779 CYPTLYQKICESELKATYRQLRGYTEDSQDLEIYLIRYNAKHETLDVPGTESVWPLSV 838
 Db 778 CYPTLYQKICESELKATYRQLRGYTEDSQDLEIYLIRYNAKHETLDVPGTESVWPLSV 837
 QY 839 ESPIGRCGEPNRCAPHEWNPDLDCSCRDGEKCAHSHHPSLDIDIGCTDLHENLGVWVV 898
 Db 838 ESPIGRCGEPNRCAPHEWNPDLDCSCRDGEKCAHSHHPSLDIDIGCTDLHENLGVWVV 897
 QY 899 FKIKTOGHARLGNLEFIEBKPLLGEALSVKAEKKWRDKREKLOLETGRVYTEAKEAV 958
 Db 898 FKIKTOGHARLGNLEFIEBKPLLGEALSVKAEKKWRDKREKLOLETGRVYTEAKEAV 957
 QY 959 DALFVDSQYNRLOADNTNIGMHAADKLVRIRAYLSELVPGVNAEIPFEELEGRIITA 1018
 Db 958 DALFVDSQYNRLOADNTNIGMHAADKLVRIRAYLSELVPGVNAEIPFEELEGRIITA 1017
 QY 1019 ISLYDARNVKGDFNNGLAQWNVKGVHDVQOQSHRSVLVPIPEWAEVSAQVPCPGGY 1078
 Db 1018 ISLYDARNVKGDFNNGLAQWNVKGVHDVQOQSHRSVLVPIPEWAEVSAQVPCPGGY 1077
 QY 1079 ILAVTAYKEGCGCVTHIEINNTDELKPKCEESEVYPTDGTNDYTAHOGTAVCN 1138
 Db 1078 ILAVTAYKEGCGCVTHIEINNTDELKPKCEESEVYPTDGTNDYTAHOGTAVCN 1137
 QY 1139 RNAGYEDAYEVDVTTASVNYKPTVEEETVTDVRDNHCEYDRGVNVPPLPAGYMTKELEY 1198
 Db 1138 RNAGYEDAYEVDVTTASVNYKPTVEEETVTDVRDNHCEYDRGVNVPPLPAGYMTKELEY 1197
 QY 1199 FPETDKWIEIGETEGKFIVDSEVLLMEE 1228

Db 1198 FPETDKWIEIGETEGKFIVDSEVLLMEE 1227
 RESULT 6
 AAR50955
 ID AAR50955 standard; protein; 1228 AA.
 XX AAR50955;
 AC AAR50955;
 DT 16-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 18-OCT-1994 (first entry)
 XX XX
 DE Bacillus thuringiensis CryIB insecticidal protoxin.
 XX XX
 KW Insecticidal crystal protein; ICP; cryIB; Ostrinea nubilalis;
 KW European corn borer; Lepidoptera; Pyralidae; toxin.
 XX XX
 OS Bacillus thuringiensis; (strain entomocidus HD 110).
 XX XX
 PN EP589110-A1.
 XX XX
 PD 30-MAR-1994.
 XX XX
 PF 19-AUG-1992; 92EP-00402307.
 XX XX
 PR 19-AUG-1992; 92EP-00402307.
 XX XX
 PA (PLBZ) PLANT GENETIC SYSTEMS NV.
 PI Peferoen M, Jansens S, Denolf P;
 DR WPI; 1994-102862/13.
 DR N-PSDB; AAQ56804.
 XX XX
 PT Method to control or combat Ostrinia - utilises Bacillus thuringiensis
 PT cry IB gene/protein for crop prevention.
 XX XX
 PS Claim 1; Page 11-18; 38pp; English.
 XX XX
 CC The CryIB toxin gene (disclosed in EP-408403) has been found to be toxic
 CC to the European corn borer (Ostrinea nubilalis). The use of CryIB
 CC insecticidal crystal protein for protecting crops against O.nubilalis is
 CC claimed. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-OCT
 CC -2003 to standardise OS field)
 XX XX
 SQ Sequence 1228 AA;
 Query Match 88.6%; Score 5740; DB 2; Length 1228;
 Best Local Similarity 89.0%; Pred.No.0;
 Matches 1097; Conservative 36; Mismatches 90; Indels 10; Gaps 3;
 QY 1 LTRNKNENEINALSIPAVSNHSTOMDLSPOARIEDSLCIAEGNNINPLVSASTVQTGI 60
 Db 1 MTSNRKNENEIN-----AVSNHSAQMDLLPDARIEDSLCIAEGNNIDPFVSASTVQTGI 55
 QY 61 NTAGRTILGVLPFFAGQIASFYSLVGLWPRGRDQWEIFLHVQOLINQOITENARNTA 120
 Db 56 NTAGRTILGVLPFFAGQIASFYSLVGLWPRGRDQWEIFLHVQOLINQOITENARNTA 115
 QY 121 LARLOGLGDSFRAYQOQSLDLEWLNDRDARTSRVLYTQYIALELDFLNAMPFLAIRNOEVP 180
 Db 116 LARLOGLGDSFRAYQOQSLDLEWLNDRDARTSRVLYTQYIALELDFLNAMPFLAIRNOEVP 175
 QY 181 LLMVYQAANLHLLLDASLFGSEFGITQEIORYYERQVEQTRDYSYCVWEYNTGLN 240
 Db 176 LLMVYQAANLHLLLDASLFGSEFGITQEIORYYERQVEQTRDYSYCVWEYNTGLN 235
 QY 241 SLRGTNAASWVRVNOFRDLTLGVLDLVALFPSPYDTRTYPINTSAQLTRVYTDALCATG 300
 Db 236 SLRGTNAASWVRVNOFRDLTLGVLDLVALFPSPYDTRTYPINTSAQLTRVYTDALCATG 295

QY 301 VNMASWYNNAPSFSAIETAVIRSPHLLDFLEQLTIFSTSSRWSATRMWYRGHTIQ 360
 Db 296 VNMASWYNNAPSFSAIAAAIRSPHLLDFLEQLTIFSASSRWSATRMWYRGHTIQ 355
 QY 361 SRPIGGGLNTSTHGSTNTSINPRLSPFSDRVYWTESYAGVLLWGIYLEPHGVPTVRFN 420
 Db 356 SRPIGGGLNTSTAGATNTSINPRLSPFSDRVYWTESYAGVLLWGIYLEPHGVPTVRFN 415
 QY 421 FRNPQNTFERGTANYSPYSPGLQKDSFELPPTTERPNYESYSHRLSHIGLISQSR 480
 Db 416 FTNPQNTSDRGATANYSPYSPGLQKDSFELPPTTERPNYESYSHRLSHIGLISQSR 475
 QY 481 VHVDPYSWTHRSADRTWTISDSITQPLVKSFNLSGTSVWSPGPTGGDIRTNVNGS 540
 Db 476 VNVFVYSWTHRSADRTWTIGENRTIQPMVKASLPQGTTVVRGPTGGDIRTNVNGS 535
 QY 541 VLSMLNFNTSLQRYVRVRYAASQTMLRVTVGGSTTFDQGPSTMSANESLTSQSR 600
 Db 536 FGPIRVTVNGPLTQRYIGFYASTVDFPVSRGGTIVNFRFLRTNWSGDELKYNFV 595
 QY 601 FAEPFVGISAGSQ-TAGISISNNAQTQTFHDKIEFIPITATFEABYDLERAQEAVAL 659
 Db 596 RRAFTPTFTQIDIIIRTSIQGLSGNGEVYIDKIEIIPVTATFEABYDLERAQEAVAL 655
 QY 660 FTNTNPRRLKTDVTDYHIDQVSNLVACLSDEFCLDEKRELLEKVKYAKRLSDERNLLQDP 719
 Db 656 FTNTNPRRLKTDVTDYHIDQVSNLVACLSDEFCLDEKRELLEKVKYAKRLSDERNLLQDP 715
 QY 720 NFTSINKQPPFISNEQSNFTSIHQSEHGHWGSENITIOGNDVFKENVYVTLPGTNEC 779
 Db 716 NFTSINKQPPFISNEQSNFTSIHQSEHGHWGSENITIOGNDVFKENVYVTLPGTNEC 775
 QY 780 YPTLYQKIGSELKAYTRYQLRGYIEDSDQLEIYLIRYNAKHETLDPGTESWPLSVE 839
 Db 776 YPTLYQKIGSELKAYTRYQLRGYIEDSDQLEIYLIRYNAKHETLDPGTESWPLSVE 835
 QY 840 SPIGRCEPNRCAPHFENPDLDCSRDGEKCAHSHHFLSDIDIGCTDLHENLGVWVWF 899
 Db 836 SPIGRCEPNRCAPHFENPDLDCSRDGEKCAHSHHFLSDIDIGCTDLHENLGVWVWF 895
 QY 900 KIKTOEGHARLGNLEFTEKPLGEALSRVRAEKKWRDKREKLOLETKRVTYAKRAVD 959
 Db 896 KIKTOEGHARLGNLEFTEKPLGEALSRVRAEKKWRDKREKLOLETKRVTYAKRAVD 955
 QY 960 ALFVDSQYRLQADTNIGMHAADKLVHRIREAYLSLSVPIGVNABIFEELEGRITAI 1019
 Db 956 ALFVDSQYRLQADTNIGMHAADKLVHRIREAYLSLSVPIGVNABIFEELEGRITAI 1015
 QY 1020 SLVDARVWVXGNDPNNGLACWNVKGVHDVQSHRSVLVPIPEWAEVVSQAVRVCPRGYI 1079
 Db 1016 SLVDARVWVXGNDPNNGLTCWNVKGVHDVQSHRSVLVPIPEWAEVVSQAVRVCPRGYI 1075
 QY 1080 LRVYAYKEGYEGCVTTHIENNTDELKFKNCEEEVYPTDTGTCNDYTAHQGTA---V 1135
 Db 1076 LRVYAYKEGYEGCVTTHIENNTDELKFKNCEEEVYPTDTGTCNDYTAHQGTA---V 1135
 QY 1136 CNSRNAGYDAYEYDVTASVNYKTYEEETVDRDNHCEYDRGYVNYPLPAGYMTKE 1195
 Db 1136 CNSRNAGYDAYEYDVTASVNYKTYEEETVDRDNHCEYDRGYVNYPLPAGYMTKE 1195
 QY 1196 LEYPPETDKWIEIGETEGKPIVDSVELLME 1228
 Db 1196 LEYPPETDKWIEIGETEGKPIVDSVELLME 1228
 RESULT 7
 ID AAY31990
 XX AAY31990 standard; protein; 1227 AA.
 AC AAY31990;
 XX
 XX 05-JAN-2000 (first entry)
 DT
 XX

DE Chimeric Cry1B insecticide HyFLiB.
 XX Insecticide; HyFLiB; crystal protein; delta-endotoxin; toxin; cry1B;
 KW cry1A(b); maize; transgenic plant; European corn borer;
 KW Ostrinia nubilalis; entomocide; crop protection; biological control.
 XX Bacillus thuringiensis.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH Protein 1..845
 FT Protein /note= "cry1B"
 FT Protein 846..1227
 FT Protein /note= "cry1A(b)"
 XX
 PN WO9950293-A1.
 XX
 PD 07-OCT-1999.
 XX
 PF 30-MAR-1999; 99WO-EP002175.
 XX
 PR 01-APR-1998; 98US-00053549.
 XX
 PA (NOVS) NOVARTIS AG.
 PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
 XX
 PI Desai NM;
 XX
 XX WPI; 1999-601323/51.
 DR N-PSDB; AAZ20086.
 XX
 FT Chimeric insecticidal protein comprising cry1B and cry1A(b) portions.
 XX
 PS Claim 12; Page 43-47; 85pp; English.
 XX
 CC This sequence represents a chimeric insecticidal protein, termed hyFLiB,
 CC that is composed of a cry1B core N-terminal toxin portion of 845 amino
 CC acids and a cry1A(b) C-terminal protoxin portion of 382 amino acids. The
 CC cry1B and cry1A(b) portions of the synthetic gene (see AAZ20086) encoding
 CC hyFLiB were derived from Bacillus thuringiensis with codon usage altered
 CC to increase expression in plants, particularly maize. When the chimeric
 CC insecticidal protein gene is expressed in transgenic maize from both PEPC
 CC and pith promoters, insecticidal activity is observed against European
 CC corn borer (*Ostrinia nubilalis*). Recombinant microbial strains
 CC transformed with the hyFLiB gene can be used in endotomocidal formulations
 CC for the biological control of Lepidopteran pests
 XX
 SQ Sequence 1227 AA;

Query Match 83.9%; Score 5436.5; DB 2; Length 1227;
 Best Local Similarity 84.3%; Pred. No. 0;
 Matches 1040; Conservative 63; Mismatches 118; Indels 13; Gaps 5;

QY 1 LTNKRKNEIINALSTPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVOTGI 60
 Db 1 MTSNRKNEIIN-----AVSNHSAQMDLSPDARIEDSLCIAEGNNINPLVSASTVOTGI 55
 QY 61 NIAGRILGVLPFPAGQIASFYSLVGLWPRGRDQWEIFLEHVEQLINQOITENARNTA 120
 Db 56 NIAGRILGVLPFPAGQIASFYSLVGLWPRGRDQWEIFLEHVEQLINQOITENARNTA 115
 QY 121 LARLQGLGDSFRAYQQSLEDWLENRDDARTSRVLYTQYIALELDPLNAMPFAIRNQEVP 180
 Db 116 LARLQGLGDSFRAYQQSLEDWLENRDDARTSRVLYTQYIALELDPLNAMPFAIRNQEVP 175
 QY 181 LLMVYQAANLHLLLDASLFGSEFGLTSQETQRYRYERQVETRDYSDYCVHYNTGLN 240
 Db 176 LLMVYQAANLHLLLDASLFGSEFGLTSQETQRYRYERQVETRDYSDYCVHYNTGLN 235
 QY 241 SLRGTNAASWVRYNQFRDITLGLVDLVALFPSPYDTRTYPINTSAQLTREVYTDAGTG 300
 Db 236 SLRGTNAASWVRYNQFRDITLGLVDLVALFPSPYDTRTYPINTSAQLTREVYTDAGTG 295

Db 479 GNTLRAPVYSWTHRSADRTNTIGPNRITQIPLVKALNLSGTVVGGFTGCDILRRTN 538
 Qy 538 NGSVLSMGLNPNNTSLQRYRVRYAASQTMVLRTVVGSTTFDQGPSPMSANESLTSQ 597
 Db 539 TGTFGDIRLNLNVLSPQRYRVRYASTTDLQFFTRINGTIVNIGNFSRTWNRGDNLEYR 598
 Qy 598 SFRFAEPVGIASGSGTAGISNNAGROTFFDKIEFIPITATFAEYDLERAQAVN 657
 Db 599 SFRTAGSTFPNFLNAOSTTTLGAQSFNSOEVIYIDRVEFPAEVTFAEYDLERAQAVN 658
 Qy 658 ALFTNTNPRLKTDTVDYHIDQVSNLVACLSDEFCLDEKRELEKVKYAKRLSDERNLLQ 717
 Db 659 ALFTSTNPRLKTDTVDYHIDQVSNLVACLSDEFCLDEKRELEKVKYAKRLSDERNLLQ 718
 Qy 718 DPNFTSINKQDPDFTSTNEQSNFTSIHQSGHGMWGSNITIOEGNDVFKENYVTLPTFN 777
 Db 719 DPNFTSINKQDPDFTSTNEQSNFTSIHQSGHGMWGSNITIOEGNDVFKENYVTLPTFN 778
 Qy 778 ECPYLYQKIGSELSKAYTRYQVYEDSODLEIYLYRYNAKHETLDVPGTESVWPLS 837
 Db 779 ECPYLYQKIGSELSKAYTRYQVYEDSODLEIYLYRYNAKHETLDVPGTESVWPLS 838
 Qy 838 VESPIGRGEPNRCAPFEWNPDLDCSRDGEKCAHSHHFSLDIDIGCTDLHENLGVW 897
 Db 839 VESPIGRGEPNRCAPFEWNPDLDCSRDGERCAHSHHFTLDDIDVGCTDLHENLGVW 898
 Qy 898 VFKIQTQEGHARLGNLFEBEKLPLGALSRVKAERKWRDKREKLOLETQVYVTEAKEA 957
 Db 899 VFKIQTQEGHARLGNLFEBEKLPLGALSRVKAERKWRDKREKLOLETQVYVTEAKEA 958
 Qy 958 VDALFVDSQYRNLQADNIGMHAADKLVHRIEAYLSSELVPGVNAEIFELEGGHII 1017
 Db 959 VDALFVDSQYRNLQADNIGMHAADKLVHRIEAYLSSELVPGVNAEIFELEGGHII 1018
 Qy 1018 AISLYDARNVVKNGDFNGLACWNVKGVDVQSHRSVLVPEWEAEVSOAVRVCPGRG 1077
 Db 1019 AMSLYDARNVVKNGDFNGLTCNVKGVDVQSHRSVLVPEWEAEVSOAVRVCPGRG 1078
 Qy 1078 YILRVTAAYKEGEGCVTIHEIENNTDELKFKNCEEEVYPTDGTGTCNDYTAHQGTAVCN 1137
 Db 1079 YILRVTAAYKEGEGCVTIHEIENNTDELKFKNCEEEVYPTDGTGTCNDYTAHQGTAVCN 1138
 Qy 1138 SRNAGVEDAYEDVTASVNTKPYEETITDVRDNHCEYDRGVNYPPLPAGVMTKELE 1197
 Db 1139 SRNAGVEDAYEDVTASVNTKPYEETITDVRDNHCEYDRGVNYPPLPAGVMTKELE 1198
 Qy 1198 YFPETDKVWEIGETEGKFIVDSVELLMEE 1228
 Db 1199 YFPETDKVWEIGETEGKFIVDSVELLMEE 1229

RESULT 9

AAW35259

 ID AAW35259 standard; protein; 1229 AA.
 AC AAW35259;

 DT 17-FEB-1998 (first entry)
 DE Bacillus thuringiensis crystal toxin CryET5.

XX EG7283; crystal toxin; CryET5; lepidopteran pest; Lymantria dispar;
 KW Ostrinia nubilalis; Pseudoplusia includens; Plutella xylostella;
 KW Spodoptera exigua; Spodoptera frugiperda; Trichoplusia ni.
 OS Bacillus thuringiensis.
 XX US5679343-A.
 PD 21-OCT-1997.
 XX 07-JUN-1995; 95US-00474038.

PR 29-JUL-1993; 93US-00100709.
 PR 30-DEC-1993; 93US-00176865.
 PA (MONS) MONSANTO CO.
 XX Jany CS, Gonzalez JM, Donovan WP, Tan Y;
 XX WPI; 1997-525682/48.
 DR N-PSDB; AAT95051.
 XX Lepidopteran toxic Bacillus thuringiensis crystal protein - useful to
 control Lepidopteran pests.
 PT Claim 2; Fig 2; 50pp; English.
 PS The present sequence is the Bacillus thuringiensis EG7283 crystal toxin
 CC CryET5, which, optionally in association with B. thuringiensis EG7283,
 CC can be used against lepidopteran pests. CryET5 is especially useful for
 CC controlling Lymantria dispar, Ostrinia nubilalis, Pseudoplusia includens,
 CC Plutella xylostella, Spodoptera exigua, Spodoptera frugiperda and
 CC Trichoplusia ni
 CC Sequence 1229 AA;
 Qy Query Match 80.8%; Score 5237.5; DB 2; Length 1229;
 Best Local Similarity 79.9%; Pred. No. 0;
 Matches 983; Conservative 94; Mismatches 149; Indels 5; Gaps 3;
 Qy 1 LTSNRKNEIINALSIPAVSNHSTOMDLSPDARIEDSLCIAEINNINPLVSASTVGTGI 60
 Db 1 LTSNRKNEIINALSIPAVSNHSTOMDLSPDARIEDSLCIAEINNINPLVSASTVGTGI 60
 Qy 61 NIAGRILGVLPFAGQIASFYSLVGEIWPGRDOWEIFLEHVEQLINQOITENARNTA 120
 Db 61 NIAGRILGVLPFAGQIASFYSLVGEIWPGRDOWEIFLEHVEQLINQOITENARNTA 120
 Qy 121 LARLOGISFRAYQOSLEDWENRRDARTRSVLYTQYIALELDLFNAMPLFAIRNOEVP 180
 Db 121 LARLOGISFRAYQOSLEDWENRRDARTRSVLYTQYIALELDLFNAMPLFAIRNOEVP 180
 Qy 181 LMVYAQAANLHLLRLDASLFGSEFGLTSOEIORYVEROVEOTRDYSDCYVWYNTGLN 240
 Db 181 LMVYAQAANLHLLRLDASLFGSEFGLTSOEIORYVEROVEOTRDYSDCYVWYNTGLN 240
 Qy 241 SLRGTNAASVRYNQFRRLDTLGLDLVALFPSSYDTRTYPINTSAQLTRVYTDIGAIG 300
 Db 241 SLRGTNAASVRYNQFRRLDTLGLDLVALFPSSYDTRTYPINTSAQLTRVYTDIGAIG 300
 Qy 301 V--NMASMWYNNNAPSPSAIETAVIRSPHLLDFLEQLTIFSTSSRSWATRHMYWRIGHT 358
 Db 301 V--NMASMWYNNNAPSPSAIETAVIRSPHLLDFLEQLTIFSTSSRSWATRHMYWRIGHT 358
 Qy 359 IQSRPIGGGLNTSTHGST--NTSINPVRLSPFSDVYVWTESYAGVLLMGVILEPIHGVPV 417
 Db 359 IQSRPIGGGLNTSTHGST--NTSINPVRLSPFSDVYVWTESYAGVLLMGVILEPIHGVPV 417
 Qy 418 RNFNRPONTFERGTANYSPQYESPGLQKXSETLPPETTERPNYESYSHRLSHIGLIS 477
 Db 418 RNFNRPONTFERGTANYSPQYESPGLQKXSETLPPETTERPNYESYSHRLSHIGLIS 477
 Qy 478 QSRVHVVPVSWTHRSADRTNTISSDITQIPLVKSFNLSNCTSVVSGPGTGGDIIRTNV 537
 Db 478 QSRVHVVPVSWTHRSADRTNTISSDITQIPLVKSFNLSNCTSVVSGPGTGGDIIRTNV 537
 Qy 538 NGSVLMSGLNFNNTSLQRYRVRYAASQTMVLRTVVGSTTFDQGPSPMSANESLTSQ 597
 Db 538 NGSVLMSGLNFNNTSLQRYRVRYAASQTMVLRTVVGSTTFDQGPSPMSANESLTSQ 597
 Qy 598 SFRFAEPVGIASGSGTAGISNNAGROTFFDKIEFIPITATFAEYDLERAQAVN 657
 Db 598 SFRFAEPVGIASGSGTAGISNNAGROTFFDKIEFIPITATFAEYDLERAQAVN 657
 Qy 658 ALFTNTNPRLKTDTVDYHIDQVSNLVACLSDEFCLDEKRELEKVKYAKRLSDERNLLQ 717

```

Db 659 ALFTSTNPRRLKTDVTDYHIDQVNMVACLSDPECLDEKRELFKVKYAKRLSDERNLQ 718
Qy 718 DPNFTSINKOPDFTSTNEQSNFTSIHQSHGWSNITIQENDVFKENYVTLPGTFN 777
Db 719 DPNFTFISGQUSFASIDQSNFFPINELSHGWSANVTIQENDVFKENYVTLPGTFN 778
Qy 778 ECPYLYLQKIGESLKAAYTRYQLRGVIEDSQDLEIYLIRYNAKHETLDVPGTGVWPLS 837
Db 779 ECPYLYLQKIGESLKAAYTRYQLRGVIEDSQDLEIYLIRYNAKHETLDVPGTGVWPLS 838
Qy 838 VESPIGRGCBPNRCAPHFENWPDLDSCRDGEKCAHSHHFSLDIDIGCTDLHENLGVV 897
Db 839 VESPIGRGCBPNRCAPHFENWPDLDSCRDGERCAHSHHFTLIDVGCTDLHENLGVV 898
Qy 898 VFKIKTOEGHARLGNLFIKPEKLLGALSRYKRAEKKWRDKREKLOLEKRVYVTEAKEA 957
Db 899 VFKIKTOEGYARLGNLEFIEKPEKLLGALSRYKRAEKKWRDKREKLOLEKRVYVTEAKEA 958
Qy 958 VDALFVDSQYNRLQADTNIGMIHAADKLVHRIEAYLSLSVIPGVNAEIPFEELEGRIT 1017
Db 959 VDALFVDSQYDQADTNIGMIHAADKLVHRIEAYLSLSVIPGVNAEIPFEELEGHIT 1018
Qy 1018 AISLYDARNVVKGDFFNGLACMNVKGHDVQOSHRSVLVPIPEWAEVSAQVRVCPGRG 1077
Db 1019 AMSLYDARNVVKGDFFNGLTCNVKKGHDVQOSHRSDLVPIPEWAEVSAQVRVCPGRG 1078
Qy 1078 YILRVTAKEGEGCVYTHIEIENNTDELKFKCEEEVYPTDGTGNDYAHQGTAVCN 1137
Db 1079 YILRVTAKEGEGCVYTHIEIENNTDELKFKCEEEVYPTDGTGNDYAHQGTAAACN 1138
Qy 1138 SRNAGYEDAYEDVTASVNYKPTVEEYTYDVRDNHCEYDRGVYVPPAGVMTKELE 1197
Db 1139 SRNAGYEDAYEDVTASVNYKPTVEEYTYDVRDNHCEYDRGVYVPPAGVMTKELE 1198
Qy 1198 YFPETDKVWIEIGTEGKFIVDSVELLMEE.1228
Db 1199 YFPETDKVWIEIGTEGKFIVDSVELLMEE.1229

```

RESULT 10

AAW17699
ID AAW17699 standard; protein; 1229 AA.

XX AAW17699;

XX 25-MAR-2003 (revised)

DT 07-JUL-1997 (first entry)

XX CryET5.

XX CryET5; cryET4; *Bacillus thuringiensis*; insecticidal crystal protein;
XX ICP; toxin; CryI protein; lepidopteran insect; insecticide.

XX *Bacillus thuringiensis*.

XX US5616319-A.

XX 01-APR-1997.

XX 30-DEC-1993; 93US-00176865.

XX 29-JUL-1993; 93US-00100709.

XX (MONS) MONSANTO CO.

XX Gonzalez JM, Donovan WP, Tan Y, Jany CS;

XX WPI; 1997-212077/19.

XX N-PSDB; AAT68434.

XX *Bacillus thuringiensis* cryET5 gene encoding insecticidal protein - useful
XX for control of lepidopteran pests.

xx
PS This sequence represents the cryET5 protein of *Bacillus thuringiensis*
XX (B.t.) isolate EG947. B.t. produces inclusions during sporulation which
CC include insecticidal crystal proteins (ICP). ICP toxins are active in
CC insects only after ingestion. Once ingested, the toxic components disrupt
CC the midgut cells, resulting in cessation of feeding, and eventually
CC death. The CryI proteins produced by B.t. are active against lepidopteran
CC insects. This protein, and the CryET4 protein (see AAW17700) belong to
CC the cryI family of ICPs. The DNA encoding this sequence can be used to
CC transform bacteria, which are useful as insecticides against a wide range
CC of lepidopteran pests, and can be applied to crops, soil and seeds. This
CC protein, or especially its toxic N terminal region, can be expressed in
CC plants, to provide protection against lepidopteran pests. The gene
CC encoding this sequence, or its fragments, can also be used to isolate
CC other similar genes. (Updated on 25-MAR-2003 to correct PF field.)
XX

Sequence 1229 AA;

Query Match 80.8%; Score 5237.5; DB 2; Length 1229;
Best Local Similarity 79.9%; Pred. NO. 0;
Matches 983; Conservative 94; Mismatches 149; Indels 5; Gaps 3;

```

Qy 1 LTSNRKNEEIIINALSIPAVSNHSTQWDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db 1 LTSNRKNEEIIINALSIPVSNPSTQWDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Qy 61 NIAGRILGVLPVPAQGIASFYSFLVGLWPRGRDQWEIFLEHVQILNQOITENARNTA 120
Db 61 NIAGRILGVLPVPAQGIASFYSFLVGLWPRGRDQWEIFLEHVQILNQOITENARNTA 120
Qy 121 LARLQGLGDSFRAYQSLDLEWLRDADTRSVLYTQYIALELDLFLNAPLFAIRNOEVP 180
Db 121 IARLEGLGRGYSYQALETWLDNRNDARSRIILERYVALELDLITAILPLFIRNEEVP 180
Qy 181 LLMVYAAQANLHLLLRDASLFGSEFGLTSQETQRYRQVEQTRDYSYCVWEYNTGLN 240
Db 181 LLMVYAAQANLHLLLRDASLFGSEGMASSDVNYQYQYQIRYTEYSNHCWQYNTGLN 240
Qy 241 SLRGTAASWVRYNQPRRLDTLGLDLVALFSDYTRTPINTSAQLTREIVYDAIGATG 300
Db 241 NLRGTNAESWLRYNQPRRLDTLGLDLVALFSDYTRTPINTSAQLTREIVYDTPGRTN 300
Qy 301 V--NMASMNWYNNNAPSFAIETAVIRSPHLLDFLEQLTIFSTSSRWSATRHMYVRGHT 358
Db 301 APSGFASWVFNWNNAPSFAIETAAIFRPHLLDFPEQLTIYSASSRWSSTQHNNYVWVGR 360
Qy 359 IQSRPIGGGLNTSTHGST-NTSINPVRLSFFSRDYYVWTSYAGVLLWGLYLEPIHGVPV 417
Db 361 LNFPRPIGGTLNTSTQGLTNNTSINPVTLQFTSRDYYVWTSNAGTNI--LFTTTPVNGVPA 418
Qy 418 RNFNRPQNTFPGTANYSQPYSPGLQKDETELPPETTERPNYESYSHRLSHLGLIS 477
Db 419 RNFNRPQNTFPGTANYSQPYSPGLQKDETELPPETTERPNYESYSHRLSHLGLII 478
Qy 478 QSRVHPVYSWTHRSADRTNTSSDSITQIPLVKSPLNLSGTSWVSGPGFTGDIIRTNV 537
Db 479 GNTLRAPVYSWTHRSADRTNTIGPNRITQIPLVKALNLHSGVTVVCGPGFTGDIIRRTN 538
Qy 538 NGSVLSMGLNFNNNTSLQRYRVRYAAQTMVLRVTVGGSTTDFDQGFPTSMANESLTSQ 597
Db 539 TGTFGDIRLNLNINPLSORVRYRVAATSTDLQFTFRINGTIVNIGNFSRTMNRGNLEVR 598
Qy 598 SRFAPFPVGISASGSGTAGISINNAGROTDFDKIEIPITATFAEYDLERAQAVN 657
Db 599 SPTAGFTFPFNLAQSTFTLGAQSFNQEVYIDRFVFPVAEVTFAEYDLERAQAVN 658
Qy 658 ALFTNTNPRRLKTDVTDYHIDQVSNLVACLSDPECLDEKRELFKVKYAKRLSDERNLQ 717
Db 659 ALFTNTNPRRLKTDVTDYHIDQVSNLVACLSDPECLDEKRELFKVKYAKRLSDERNLQ 718
Qy 718 DPNFTSINKOPDFTSTNEQSNFTSIHQSHGWSNITIQENDVFKENYVTLPGTFN 777

```

```

Db 719 DPVFTFISQSLFASIDQCNFNSINLSHGWSANVTIQGNDVFKENYVTLPGTFN 778
Qy 778 ECPYTYLYQKIGSELKAYTRYQLRGYIEDSQBLEIYLIRYNAKHETLDVPGTESVMPLS 837
Db 779 ECPYTYLYQKIGSELKAYTRYQLRGYIEDSQBLEIYLIRYNAKHETLDVPGTDSLWPLS 838
Qy 838 VESPIGRGPNRCAPHFENPDLDCSCRGCEKCAHSHHSFSLDIDIGCTDLHENLGVWV 897
Db 839 VESPIGRGPNRCAPHFENPDLDCSCRDGERCAHSHHFTLDDVGCCTDLHENLGVWV 898
Qy 898 VFPIKTOEGHARLGNLEFIEKPLIGALSQRVKAEEKWDRKKLQLETKRVVYTAKEA 957
Db 899 VFPIKTOEGHARLGNLEFIEKPLIGALSQRVKAEEKWDRKKLQLETKRVVYTAKEA 958
Qy 958 VDALFVDSQYNRLQADTNIGMHAADKLVRIRIERYLSELVIPGVNAEIEFELEGRIT 1017
Db 959 VDALFVDSQYDQLQADTNIGMHAADKLVRIRIERYLSELVIPGVNAEIEFELEGHIT 1018
Qy 1018 AISLYDARNVVKNGDFNGLACNVKGVHDVQSHHRSVLVIVPEWAEVSGAVRVCPGRG 1077
Db 1019 AMSLYDARNVVKNGDFNGLTCNVKGVHDVQSHHRSDLVIVPEWAEVSGAVRVCPGRG 1078
Qy 1078 YILRVYAYKEGYGECVYTHIEINNTDELKFKNCEEEVYPTDTGTCNDYTAHOGTAVCN 1137
Db 1079 YILRVYAYKEGYGECVYTHIEINNTDELKFKNCEEEVYPTDTGTCNDYTAHOGTAACN 1138
Qy 1138 SRNAGVEDAYEVDVTASVNYKPYEETVYDVRDNHCEYDRGVNYPPLPAGVMTKELE 1197
Db 1139 SRNAGVEDAYEVDVTASVNYKPYEETVYDVRDNHCEYDRGVNYPVPAGVYVKELE 1198
Qy 1198 YFPETDKVWIEIGETEGKFIVDSVELLLMEE 1228
Db 1199 YFPETDKVWIEIGETEGKFIVDSVELLLMEE 1229

```

RESULT 11

```

AAW87633
ID AAW87633 standard; protein; 1229 AA.
AC AAW87633;
XX
XX 03-MAR-1999 (first entry)
XX
XX CryET5 protein sequence.
XX
XX CryET4; Bacillus thuringiensis strain EG5847; crystal toxin; CryET5;
XX Lepidoptera; Helicoverpa zea; resistant strain; Plutella xylostella;
XX Spodoptera exigua; S. frugiperda; S. frugiperda; Trichoplusia ni.
XX
XX Bacillus thuringiensis.
XX
XX US5854053-A.
XX
XX 29-DEC-1998.
XX
XX 06-JAN-1997; 97US-00779046.
XX
XX 29-JUL-1993; 93US-00100709.
XX 30-DEC-1993; 93US-00176865.
XX 07-JUN-1995; 95US-00474038.
XX
XX (ECOG-) ECOGEN INC.
XX
XX Donovan WP, Gonzalez JM;
XX
XX WPI; 1999-094915/08.
XX DR N-PSDB; AAW83927.
XX
XX New strains EG5847 and EG10368 of Bacillus thuringiensis - producing
XX crystal toxins active against Lepidoptera, particularly Helicoverpa zea,
XX Plutella xylostella and spodoptera species.
XX

```

PS Example 2; Fig 2A-J; 49pp; English.

XX The present sequence represents a CryET5 protein. The protein is isolated
 CC from a new strain of Bacillus thuringiensis (B.t.), strain EG5847 (NRRL B
 CC -21110). EG5847 contains two new crystal toxin genes (cryET4 and cryET5)
 CC that produce proteins active against Lepidoptera. CryET4 is more active
 CC than CryIA(a) against Helicoverpa zea resistant-strains of Plutella
 CC xylostella, Spodoptera exigua and S. frugiperda, while CryET5 is more
 CC active than CryIB against S. frugiperda and Trichoplusia ni (and both are
 CC active against additional insect species)

XX Sequence 1229 AA;

Query Match 80.8%; Score 5237.5; DB 2; Length 1229;
 Best Local Similarity 79.9%; Pred. No. 0;
 Matches 983; Conservative 94; Mismatches 149; Indels 5; Gaps 3;

```

Qy 1 LTSNRKNEININALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVOTGI 60
Db 1 LTSNRKNEININALSIPVSNSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVOTGI 60
Qy 61 NIAGRILGVLPFAQGIASFYSFLVGLMPGRDQWEIFLEHVEQLINQOITENARNTA 120
Db 61 NIAGRILGVLPFAQGIASFYSFLVGLMPGRDQWEIFLEHVEQLIRQOITENTNTA 120
Qy 121 LARLQGLGDSFRAYQOSLEDWLENRDDARTSRVLYTQYIALELDFLNMPFLAIRNOEVP 180
Db 121 LARLQGLGDSFRAYQOSLEDWLENRDDARTSRVLYTQYIALELDFLNMPFLAIRNOEVP 180
Qy 181 LLMVYQAANLHLLLRDASLFGSEFGLTSQETQRYVYEROVEOTRDYDVCVWYNTGLN 240
Db 181 LLMVYQAANLHLLLRDASLFGSEFGLTSQETQRYVYEROVEOTRDYDVCVWYNTGLN 240
Qy 241 SLRGTNAASVRYNQFRDRLTLGVLDLVALFPSTYDTRTYPIINTSAQLTREYVYDAIGATG 300
Db 241 SLRGTNAASVRYNQFRDRLTLGVLDLVALFPSTYDTRTYPIINTSAQLTREYVYDAIGATG 300
Qy 301 V--NMASMMWYNNNAPSFSAIETAVIRSPHLLDFLEQLTIFSTSSRWASATRHMTYWRGHT 358
Db 301 APSGFASNTMFWNNAPSFSAIETAAIIRPPHLLDFPEQLTIYSASSRWSSQHMNYWYGH 360
Qy 359 IQSRPIGGGLNTSGTSTNTSINPVRLSPFSRDVYVWTSYAGVLLWGIYLEPIHGVPTV 417
Db 361 LNFRPIGGTLNTSQGLTNTSINPVTLQTSRDVYRTESNAGTNI--LFTTPVNGVPWA 418
Qy 418 RNFNRPNQTFERTGANYSQYSPGLQKDSLETLPETTERPNYSESYSHRLSHIGLIS 477
Db 419 RNFNFINQIYERGATTSQYQGVGLQFDSETLPETTERPNYSESYSHRLSHIGLII 478
Qy 478 QSRVHVYVSWTHRSADRTNITSSDSITQIPLVKSFNLSGTSVSGPGTGGDIIRTNV 537
Db 479 GNTLRAPVYSWTHRSADRTNITGPNRITQIPLVKALNLSHSGVTVGPGTGGDILARTN 538
Qy 538 NGSVLSGLNFINNTSLQRYRVRYAAASQTMVLRTVTVGSGSTTFDQGPSPMSANESLTSQ 597
Db 539 TGTFGDIRLNIINVPVLSQRYRVRYRYASTTDQFTFRINGTIVNIGNFSRTNRGNDNLEYR 598
Qy 598 SFRFAEPFVPGISASGSGQTAGISISNNAGRTQTHFDKIEFIPITATFEAYDLERAQAVN 657
Db 599 SFRTAGFSTPPFNLAQSTFTLGAQSFNSQEVVIDRVEFVPAEVTFEAYDLERAQAVN 658
Qy 658 ALFTNTNPRRLKTDVTDYHIDQVSNLVACLSDSFCLDEKRELLEKVKYAKRLSDERNLLQ 717
Db 659 ALFTSTNPRRLKTDVTDYHIDQVSNLVACLSDSFCLDEKRELPEKVKYAKRLSDERNLLQ 718
Qy 718 DPNFTSINKOPDFISTNEQSNFTSIHQSEHGWSGSENITIOEGNDVFKENYVTLPGTFN 777
Db 719 DPNFTFISQSLFASIDGOSNFFSINLSHGWSANVTIQGNDVFKENYVTLPGTFN 778
Qy 778 ECPYTYLYQKIGSELKAYTRYQLRGYIEDSQBLEIYLIRYNAKHETLDVPGTESVMPLS 837
Db 779 ECPYTYLYQKIGSELKAYTRYQLRGYIEDSQBLEIYLIRYNAKHETLDVPGTDSLWPLS 838

```


QY 838 VESPIGRCEPNRCAPHEWNPDLDCSDGCKEKAHHSHFSLDIDIGCTDLHENLGVWV 897
 Db 839 VESPIGRCEPNRCAPHEWNPDLDCSDGCKEKAHHSHFSLDIDIGCTDLHENLGVWV 898
 QY 898 VFKIQTQGHARLGNLEFIEEKPLIGEALSRVGAERKWRDKREKLOLETRVYTEAKEA 957
 Db 899 VFKIQTQGHARLGNLEFIEEKPLIGEALSRVGAERKWRDKREKLOLETRVYTEAKEA 958
 QY 958 VDALFVDSQYNRLQADTNIGMHAADKLVRIRIAYLSLSVIFGVNAEIEELEGRIIT 1017
 Db 959 VDALFVDSQYNRLQADTNIGMHAADKLVRIRIAYLSLSVIFGVNAEIEELEGRIIT 1018
 QY 1018 AISLYDARNVVKNGDFNGLACWNVKGVHDVQOQSHHRSVLVPEWEAEVSQAVRVCGRG 1077
 Db 1019 AMSLYDARNVVKNGDFNGLACWNVKGVHDVQOQSHHRSVLVPEWEAEVSQAVRVCGRG 1078
 QY 1078 YILRVYAYKEGEGCVTIHIEINNTDELKPKNCEEEVYPTDGTGNDYTAHQTAACN 1137
 Db 1079 YILRVYAYKEGEGCVTIHIEINNTDELKPKNCEEEVYPTDGTGNDYTAHQTAACN 1138
 QY 1138 SRNAGYEDAYEVDVTTASVNYKPTVEEETDVRDNHCEYDRGVYVPPYPAGVYTKELE 1197
 Db 1139 SRNAGYEDAYEVDVTTASVNYKPTVEEETDVRDNHCEYDRGVYVPPYPAGVYTKELE 1198
 QY 1198 YFPETDKVWIEBIGTEGKFIIVDSVELLMEE 1228
 Db 1199 YFPETDKVWIEBIGTEGKFIIVDSVELLMEE 1229

RESULT 12

AY30923
 ID AY30923 standard; protein; 1229 AA.
 AC AY30923;

DT 18-OCT-1999 (first entry)

XX B. thuringiensis cryET5 partial protein.

XX Toxin; cryET4; cryET5; insecticidal; Lepidoptera; transformed plant;
 KW crystal protein; insect.

XX Bacillus thuringiensis.

XX Key Location/Qualifiers

FT Protein 1. 1229
 FT /note= "partial protein sequence. No ATG start codon
 given"

XX US5942658-A.

XX 24-AUG-1999.

XX 24-JUN-1997; 97US-00881340.

XX 29-JUL-1993; 93US-00100709.

XX 30-DEC-1993; 93US-00176865.

XX 07-JUN-1995; 95US-00474038.

XX (MONS) MONSANTO CO.

XX Gonzalez JM, Jany CS, Tan Y, Donovan WP;

XX WPI; 1999-493544/41.

XX N-PSDB; AAZ09160.

XX Transformed plant comprising insecticidal crystal proteins.

XX Claim 3; Fig 2A-J; 50pp; English.

XX This invention describes novel transformed plants containing Bacillus

XX thuringiensis strain EG5847 insecticidal crystal protein genes cryET4 or

XX cryET5. cryET4 and cryET5 are novel toxin genes which produce

CC insecticidal proteins with activity against a broad spectrum of insects
 CC of the order Lepidoptera. This sequence represents the cryET5 protein
 CC described in the method of the invention
 XX
 SQ Sequence 1229 AA;

Query Match 80.8%; Score 5237.5; DB 2; Length 1229;
 Best Local Similarity 79.9%; Pred. No. 0;
 Matches 983; Conservative 94; Mismatches 149; Indels 5; Gaps 3;

QY 1 LTSNRKNEIINALSIPAVSNEHSTOMDLPDARIEDSLCIAEGNNINPLVSASTVOTGI 60
 Db 1 LTSNRKNEIINALSIPAVSNEHSTOMDLPDARIEDSLCIAEGNNINPLVSASTVOTGI 60
 QY 61 NIAGRILGVGVFPAGQIASFYSLVGLMWRGRDOWEIFLEHVQOLINOITENARNTA 120
 Db 61 NIAGRILGVGVFPAGQIASFYSLVGLMWRGRDOWEIFLEHVQOLINOITENARNTA 120
 QY 121 LARLQGLGDSFRAYQOSLEDWLENRDDARTSRVLYTCYITALELDFLNAMEFLAIRQOEVP 180
 Db 121 LARLQGLGDSFRAYQOSLEDWLENRDDARTSRVLYTCYITALELDFLNAMEFLAIRQOEVP 180
 QY 181 LLMVYAQAANLHLLLRDASLFGSEFGLTSQEIQRYERQVEQTRDYSDYCVWEYNTGLN 240
 Db 181 LLMVYAQAANLHLLLRDASLFGSEFGLTSQEIQRYERQVEQTRDYSDYCVWEYNTGLN 240
 QY 241 SLRGTAASWVRYNQFRDLTGLVDLVALFPSYDTRTYPIINTSAQLTREYVTDATGATG 300
 Db 241 SLRGTAASWVRYNQFRDLTGLVDLVALFPSYDTRTYPIINTSAQLTREYVTDATGATG 300
 QY 301 V--NMASMNWYNNNAPSFAIETAVIRSHLLDPLQLATIFSTSSRWASATRHMYTNRGHT 358
 Db 301 V--NMASMNWYNNNAPSFAIETAVIRSHLLDPLQLATIFSTSSRWASATRHMYTNRGHT 358
 QY 359 IQSRPIGGGLNTSTHGST-NTSINPVLRSFFSRDVTYTESYAGVLLMWLYLEPIHGVTV 417
 Db 359 IQSRPIGGGLNTSTHGST-NTSINPVLRSFFSRDVTYTESYAGVLLMWLYLEPIHGVTV 417
 QY 418 RNFNPQNTFERGTANYQPYESPGQLKDETELPPETTERPNYTESYSHRLSHGLIS 477
 Db 418 RNFNPQNTFERGTANYQPYESPGQLKDETELPPETTERPNYTESYSHRLSHGLIS 477
 QY 478 QSRVHVYVSWTHRSADRTNTISSDSITQIPVKSFNLSNGTSVVSQPGFTGGDIIRTNV 537
 Db 478 QSRVHVYVSWTHRSADRTNTISSDSITQIPVKSFNLSNGTSVVSQPGFTGGDIIRTNV 537
 QY 479 GNTLRAPVYVSWTHRSADRTNTIGPNRIQTQIPVKSFNLSNGTSVVSQPGFTGGDIIRTNV 538
 Db 479 GNTLRAPVYVSWTHRSADRTNTIGPNRIQTQIPVKSFNLSNGTSVVSQPGFTGGDIIRTNV 538
 QY 538 NGSVLSMGLNFNTSLQRYRVRYAASQTMVLRVVGGSTTFDQGPFTMSANESLTSQ 597
 Db 538 NGSVLSMGLNFNTSLQRYRVRYAASQTMVLRVVGGSTTFDQGPFTMSANESLTSQ 597
 QY 598 SPFRFAEPVVGISAGSOTAGISINNAGQTFHFDKIEFIPITATPEAEVDLERACEAVN 657
 Db 598 SPFRFAEPVVGISAGSOTAGISINNAGQTFHFDKIEFIPITATPEAEVDLERACEAVN 657
 QY 658 ALFTNTNPRRLKTDVTDYHIDQVSNLVACLSEDFCLDEKRELEKVKYAKRLSDERNLQ 717
 Db 658 ALFTNTNPRRLKTDVTDYHIDQVSNLVACLSEDFCLDEKRELEKVKYAKRLSDERNLQ 717
 QY 718 DPNFTSINKQDPFISTNEQSNFTSHEQSEHGWGSENITIQENDVPKENVYTLPGTGN 777
 Db 718 DPNFTSINKQDPFISTNEQSNFTSHEQSEHGWGSENITIQENDVPKENVYTLPGTGN 777
 QY 778 ECVPTYLYQKIGESSELKAYTRYQLRGYIEDSQLEIYLIRYNAKHETLDPVGTESVWPLS 837
 Db 778 ECVPTYLYQKIGESSELKAYTRYQLRGYIEDSQLEIYLIRYNAKHETLDPVGTESVWPLS 837
 QY 838 VESPIGRCEPNRCAPHEWNPDLDCSDGCKEKAHHSHFSLDIDIGCTDLHENLGVWV 897
 Db 838 VESPIGRCEPNRCAPHEWNPDLDCSDGCKEKAHHSHFSLDIDIGCTDLHENLGVWV 898
 QY 898 VFKIQTQGHARLGNLEFIEEKPLIGEALSRVGAERKWRDKREKLOLETRVYTEAKEA 957
 Db 898 VFKIQTQGHARLGNLEFIEEKPLIGEALSRVGAERKWRDKREKLOLETRVYTEAKEA 958

Qy 958 VDALFVDSQYNRLQADTNIGMHAADKLVHRIIRAYLSLSVIPGVNAEIPFEELEGRIT 1017
 Db 959 VDALFVDSQYDQLQADTNIGMHAADKLVHRIIRAYLSLSVIPGVNAEIPFEELEGRIT 1018
 Qy 1018 AISLYDARNVVKNGDFNGLACNNVKGHDVQOQSHRSVLVPEWEAEVSQAVRVCPRG 1077
 Db 1019 AMSLYDARNVVKNGDFNGLACNNVKGHDVQOQSHRSVLVPEWEAEVSQAVRVCPRG 1078
 Qy 1078 YILRVTAKEGEGCGCVTIHEIENNTDELKFKNCEEEVPTDGTGTCNDYTAHQGTAVCN 1137
 Db 1079 YILRVTAKEGEGCGCVTIHEIENNTDELKFKNCEEEVPTDGTGTCNDYTAHQGTAACN 1138
 Qy 1138 SRNAGEDAYEVDVTASVNVKPYEETTYDVRDNCHEVDGRGVNTPPLPAGVMTKELE 1197
 Db 1139 SRNAGEDAYEVDVTASVNVKPYEETTYDVRDNCHEVDGRGVNTPPVPGVYVTKELE 1198
 Qy 1198 YFPETDKVWIEIGTEGKFVDSVELLMEE 1228
 Db 1199 YFPETDKVWIEIGTEGKFVDSVELLMEE 1229

RESULT 13

ADK98479
 ID ADK98479 standard; protein; 1229 AA.

AC ADK98479;
 XX
 DT 03-JUN-2004 (first entry)
 XX

DE B thuringiensis cry1Bb partial protein sequence SeqID2.
 KW insecticidal protein; plant; pesticide; gene therapy;
 KW lepidopteran insect pest; transgenic plant;
 KW insect infestation resistance; monocot; dicot; cry1Bb.
 XX
 OS Bacillus thuringiensis.

XX WO2004020636-A1.
 XX 11-MAR-2004.
 XX 26-AUG-2003; 2003WO-US026510.
 XX 29-AUG-2002; 2002US-0407428P.

XX (MONS) MONSANTO TECHNOLOGY LLC.
 PA (BOGD/) BOGDANOVA N N.
 PA (ROMA/) ROMANO C P.

XX Bogdanova NN, Romano CP;
 XX WPI: 2004-269221/25.
 XX N-PSDB; ADK98478.

XX New polynucleotide sequence optimized for expression of an insecticidal protein in a plant, useful in the control of lepidoptera insect pests, and for producing transgenic plants with the ability to resist insect infestations.

XX Claim 7; SEQ ID NO 2; 138pp; English.

XX This invention relates to a novel polynucleotide sequence optimised for expression of an insecticidal protein in a plant. The invention may be useful for the production of pesticides whilst the disclosed sequences may be used for gene therapy. The polynucleotide sequence and methods are useful in the control of lepidopteran insect pests, and for producing transgenic plants with the ability to resist insect infestations. The invention provides polynucleotide sequences with enhanced, improved and optimised expression in monocot and dicot plant species. The present sequence is that of the (partial) B thuringiensis cry1Bb protein which is related to the invention.

XX

SQ Sequence 1229 AA;

Query Match 80.8%; Score 5237.5; DB 8; Length 1229;
 Best Local Similarity 79.9%; Pred. No. 0;
 Matches 983; Conservative 94; Mismatches 149; Indels 5; Gaps 3;

Qy 1 LTSNRKNEIINALSTPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVGTGI 60
 Db 1 LTSNRKNEIINALSIPTVSNSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVGTGI 60
 Qy 61 NIAGRILGLVGVFAGQIASFYSLFVGLWPRGRDOWEIEFLEHVQELINOQITENANNTA 120
 Db 61 NIAGRILGLVGVFAGQIASFYSLFVGLWPRGRDOWEIEFLEHVQELINOQITENANNTA 120
 Qy 121 LARLOGLGDSFRAYQOQSLDMLNRRDARTSRVLYTOYIALELDLFLNAPLFAIRNOEVP 180
 Db 121 IARLEGLGRGYSYQQALETWLDNRNDARSITILERYVALELDITTAIPLFIRNRSEVP 180
 Qy 181 LLMVYQAANLHLLLRDASLFGSEFGLTSQETQRYRQVEQTRDYSDYCVENYNTGLN 240
 Db 181 LLMVYQAANLHLLLRDASLFGSEGMASDVNQYQEQIRYTEESNHCVOYNTGLN 240
 Qy 241 SLRGTNAASWVRYNQFRRDLTLGVLDLVALFPSPVDTRTYPINTSAQLTRVYTDAGATG 300
 Db 241 NLRGTAESWLRYNQFRRDLTLGVLDLVALFPSPVDTRTYPINTSAQLTRVYTDAGATG 300
 Qy 301 V--NMASMMYNNNAPSFAIETAVIRSPHLLDPLEQLTIFSTSSRSWATRHMTYWRGHT 358
 Db 301 APSGFASSTWNNNAPSFAIETAAIPRPHLLDPLEQLTIFSTSSRSWATRHMTYWRGHT 360
 Qy 359 IQSRPIGGGLNTSTHGST-NTSINPVRLSFFSDVYVWTESYAGVLLGWILEPHIGVPTV 417
 Db 361 LNFRPIGGTTLNTSTQGLTNNTSINPVTLQFTSDVYVWTESNAGTNI--LFTTPVNGVPWA 418
 Qy 418 RFNFRPQNTFERGTANYOPVESPGLOKDSLTELPPETTERPNYESYSHRLSHIGLIS 477
 Db 419 RFNFRPQNTFERGTANYOPVESPGLOKDSLTELPPETTERPNYESYSHRLSHIGLII 478
 Qy 478 QSRVHVSVSWTHRSADRTNTISSDSITQIPLVKSNLNSGTSVVSQPGTGGDIIRTNV 537
 Db 479 GNTLRAPVSVSWTHRSADRTNTIGPNRITQIPLVKALNLHSGVTVVGPGTGGDILARTN 538
 Qy 538 NGSVLGMNLFNNTSLQRYRVRYAASQTMVLRVTVGGSTTDFDQGPSTMSANESLTSQ 597
 Db 539 TGTFGDIRLNLNINPLSQRVRYRYASTTDLQFTFRINGTVAIGNFSRTMRGDNLEYR 598
 Qy 598 SRFPAEFPVGISASGSGTAGISISNNAGROTFFDKIEFIPITATFEAYDLERAQAVN 657
 Db 599 SRFAGFSTPFNFLNAQSTFTLGAQSFNSQEVYIDRVEFPAEVTFEAYDLERAQAVN 658
 Qy 658 ALFTNTNPRRLKTDVTDYHIDQVSNLVACLSDFCLDEKRELEKVKYAKRLSDERNLLQ 717
 Db 659 ALFTSINPRRLKTDVTDYHIDQVSNLVACLSDFCLDEKRELEKVKYAKRLSDERNLLQ 718
 Qy 718 DPNFTSINKQDPFISNTEQSNFTSIHQSHGWGSENITIQGNDVFKENYVTLPGTFN 777
 Db 719 DPNFTSINKQDPFISNTEQSNFTSIHQSHGWGSENITIQGNDVFKENYVTLPGTFN 778
 Qy 778 ECPYTVLYOKIGESSELKAYTRYQLRGVIEDSQLEIYLIIRYNAKHETLDVPGTESVPLS 837
 Db 779 ECPYTVLYOKIGESSELKAYTRYQLRGVIEDSQLEIYLIIRYNAKHETLDVPGTESVPLS 838
 Qy 838 VESPIGRCEPNRCAPHFENWPDLCSCRDGECACAHSHHFLSDIDTIGCTDLHENLGWV 897
 Db 839 VESPIGRCEPNRCAPHFENWPDLCSCRDGECACAHSHHFLSDIDTIGCTDLHENLGWV 898
 Qy 898 VFYKIQEGHARLGNLEFIEEKPLGELASRVKRAEKKWRDKREKLOLETYKRVYTEAKEA 957
 Db 899 VFYKIQEGHARLGNLEFIEEKPLGELASRVKRAEKKWRDKREKLOLETYKRVYTEAKEA 958
 Qy 958 VDALFVDSQYNRLQADTNIGMHAADKLVHRIIRAYLSLSVIPGVNAEIPFEELEGRIT 1017
 Db 959 VDALFVDSQYDQLQADTNIGMHAADKLVHRIIRAYLSLSVIPGVNAEIPFEELEGRIT 1018

QY 1018 AISLYDARNVVKNGDFNNGLACMNKGVHVDVQOQSHRSVLVPEWEAEVSQAVRVCPRG 1077
 DB 1019 AMSLYDARNVVKNGDFNNGLTCNWKGVHVDVQOQSHRSVLVPEWEAEVSQAVRVCPRG 1078
 QY 1078 YILRVATYKEGYGEGCVTHIEINNTDELKPKNCEEEVPTDGTGTCNDYTAHQGTAVCN 1137
 DB 1079 YILRVATYKEGYGEGCVTHIEINNTDELKPKNCEEEVPTDGTGTCNDYTAHQGTAAACN 1138
 QY 1138 SRNAGYEDAVEVDVTASVNYKPTVEEETVTDVDRDNHCEVDRGVNYPPLPAGVWYKELE 1197
 DB 1139 SRNAGYEDAVEVDVTASVNYKPTVEEETVTDVDRDNHCEVDRGVNYPVPAGVYVTKELE 1198
 QY 1198 YFPETDKVWIEIGTEGKFIIVDSVELLMEE 1228
 DB 1199 YFPETDVWIEIGTEGKFIIVDSVELLMEE 1229

RESULT 14
 ADK98484
 ID ADK98484 standard; protein; 1230 AA.
 XX AC ADK98484;
 XX 03-JUN-2004 (first entry)
 DE B thuringiensis cryIbB-related expression cassette protein SeqID7.
 XX insecticidal protein; plant; pesticide; gene therapy;
 KW lepidopteran insect pest; transgenic plant;
 KW insect infestation resistance; monocot; dicot; cryIbB.
 XX Bacillus thuringiensis.
 OS Synthetic.
 OS WO2004020636-A1.
 XX 11-MAR-2004.
 XX 26-AUG-2003; 2003WO-US026510.
 XX 29-AUG-2002; 2002US-0407428P.
 XX (MONS) MONSANTO TECHNOLOGY LLC.
 PA (BOGD/) BOGDANOVA N N.
 PA (ROMA/) ROMANO C P.
 XX Bogdanova NN, Romano CP;
 PI WPI; 2004-269221/25.
 DR N-PSDB; ADK98482.

XX New polynucleotide sequence optimized for expression of an insecticidal
 PT protein in a plant, useful in the control of Lepidoptera insect pests,
 PT and for producing transgenic plants with the ability to resist insect
 PT infestations.
 XX Claim 7; SEQ ID NO 7; 138pp; English.
 XX This invention relates to a novel polynucleotide sequence optimised for
 CC expression of an insecticidal protein in a plant. The invention may be
 CC useful for the production of pesticides whilst the disclosed sequences
 CC may be used for gene therapy. The polynucleotide sequence and methods are
 CC useful in the control of lepidopteran insect pests, and for producing
 CC transgenic plants with the ability to resist insect infestations. The
 CC invention provides polynucleotide sequences with enhanced, improved and
 CC optimised expression in monocot and dicot plant species. The present
 CC sequence is that of a B thuringiensis cryIbB expression cassette protein
 CC which is related to the invention.
 XX Sequence 1230 AA;

Query Match 80.8%; Score 5233.5; DB 8; Length 1230;

Best Local Similarity 79.8%; Pred. No. 0;
 Matches 982; Conservative 94; Mismatches 149; Indels 5; Gaps 3;
 QY 2 TSNRKNENEIINALSIPAVSNHSTQMDLSPDARIEDSLCTAEGNNINPLVASTVQTGIN 61
 DB 3 TSNRKNENEIINALSIPAVSNHSTQMDLSPDARIEDSLCTAEGNNINPLVASTVQTGIN 62
 QY 62 IACRIILGVLPAGQIASFYSPLVGLMPRGDOWEIEFLEHVEQLINQOITENARTAL 121
 DB 63 IACRIILGVLPAGQIASFYSPLVGLMPRGDOWEIEFLEHVEQLINQOITENARTAL 122
 QY 122 ARLOGLGDSFRAYQOQSLDLENRDDRARTSVLYTQYIALLELDFLAMPFLAIRNOEVL 181
 DB 123 ARLEGLGRGYSYQOALETWLDNRNDARSRIILERYVALELDITTAIPLFRIRNEEVL 182
 QY 182 LMVYQAANLHLLLRDASLFGSEFGLTQEIQYRYERQVEQTRDYSDYCEVWYTNGLNS 241
 DB 183 LMVYQAANLHLLLRDASLFGSEMGWASDVNQYQEIQIRYTEEYSHNCVOWYNTGLNN 242
 QY 242 LRGITNAASWVRVNOFRDLTLGVLDLVALFPSYDTRTYPINTSAQLTREYVTDATGATV 301
 DB 243 LRGITNAESWLRVNOFRDLTLGVLDLVALFPSYDTRTYPINTSAQLTREYVTDATGATV 302
 QY 302 --NMASMNWYNNNAPSFAIETAVIRSPHLLDFLEQLTIFSTSSRWSATRHMTYWRGHTI 359
 DB 303 PSGFASTNFWNNAPSFAIEAAIPRPHLLDFPEQLTIYSASSRWSSTQHMYWVGHRL 362
 QY 360 QSRPIGGGLNTSTHGST-NYSINPVLSPFRSDRVYWTESYAGVLLMGIVLEPHGVPTVR 418
 DB 363 NFRPIGGTLNTSTQGLTNTNTSINPVLQFTSRDVRVYTESNAGTNI--LFTTPVGPVWAR 420
 QY 419 FNRPNQNTFERGTANYSQPYSPGLQKDSLETLPETTERPNVYESYSHRSHIGLISQ 478
 DB 421 FNFINQNIYERGATYSQPYQGVGLQDFSETLPETTERPNVYESYSHRSHIGLIIIG 480
 QY 479 SRVHPVYSWTHRSADRTNTISSDSITQIPLVKSFNLNSGTSVVSGPGTGGDIIRTNVN 538
 DB 481 NTLRAPVYSWTHRSADRTNTIGPNRITQIPLKALNLSHGVTVVGGPGTGGDILARTNT 540
 QY 539 GSVLSMGLNPNNTSLQRYVRVRYAASOTWLVRTVVGSGTTFDQGPSPMSANESITSS 598
 DB 541 GTFPGIRLNVPLSQRVRYRYASTTDLQFTRINGTIVNIGNFSRTNRNNDNLEYS 600
 QY 599 FRFAEPFVGISASGSGTAGISISNNAAGROTFFDKIEFIPTATFEAEYDLERAQAVNA 658
 DB 601 FRTAGFTFPFNLAQSTFTLGAQSFNSQEVYIDRVFVPAEVTFEAEYDLERAQAVNA 660
 QY 659 LFTNTNPRRLKTDVTDYHIDQVSNLVACLSDFCLDEKRELEKVKYAKRLSDERNLLQD 718
 DB 661 LFTSTNPRRLKTDVTDYHIDQVSNLVACLSDFCLDEKRELEKVKYAKRLSDERNLLQD 720
 QY 719 PNFTSINKOPDFTSTNEQSNFTSIHQSEHGWGSENITIQEGNDVFKENYVTLPGTFNE 778
 DB 721 PNFTFISGQSLSFASIDQSNFFSINELSHGWGSENITIQEGNDVFKENYVTLPGTFNE 780
 QY 779 CYPTVLYQKIGESLKAITYRQYLGVIYEDSQLEIYIRYNAKHETLDVPGTESVWPLSV 838
 DB 781 CYPTVLYQKIGESLKAITYRQYLGVIYEDSQLEIYIRYNAKHETLDVPGTESVWPLSV 840
 QY 839 ESPIGRCGPNRCAPHFENWPDLCSCRDGECAGHSHHFSLDIDIGCTDLHENLGVWV 898
 DB 841 ESPIGRCGPNRCAPHFENWPDLCSCRDGECAGHSHHFTLIDVCGTDLHENLGVWV 900
 QY 899 FKIKTOEGHARLGNLEFIEEKPLGREALSRVRAEKKWRDEKREKLETKRVVTEAKEAV 958
 DB 901 FKIKTOEGHARLGNLEFIEEKPLGREALSRVRAEKKWRDEKREKLETKRVVTEAKEAV 960
 QY 959 DALFVDSQYNRLQADTNIGMIIHAADKLVRIRIYALSELVPIGVNAEIFEELGRIITA 1018
 DB 961 DALFVDSQYDQLQADTNIGMIIHAADKLVRIRIYALSELVPIGVNAEIFEELGRIITA 1020
 QY 1019 ISLYDARNVVKNGDFNNGLACMNKGVHVDVQOQSHRSVLVPEWEAEVSQAVRVCPRG 1078

Db 1021 MSLYDARNVVKNGDFNGLTCWNVKGHDVQQSHRSDLVIPWEAEVSQAVRVCPRGY 1080
 Qy 1079 ILRVTAKEGEGECVTIHEIENNTDELKFKNCBEEVYPTDCTCNDYTAHQGTAVCNS 1138
 Db 1081 ILRVTAKEGEGECVTIHEIENNTDELKFKNCBEEVYPTDCTCNDYTAHQGTAAACNS 1140
 Qy 1139 RNAGYEDAYEDVTASVNYKPTVEEYTYDVRDNHCEYDGVYVYPLPAGYMTKELEY 1198
 Db 1141 RNAGYEDAYEDVTASVNYKPTVEEYTYDVRDNHCEYDGVYVYVPPVAGYVTKLEY 1200
 Qy 1199 FPKTDKWIIEIGETEGKFIIVDSVELLMEE 1228
 Db 1201 FPKTDKWIIEIGETEGKFIIVDSVELLMEE 1230

RESULT 15
 ADK98489
 ID ADK98489 standard; protein; 1230 AA.
 XX
 AC ADK98489;
 DT 03-JUN-2004 (first entry)
 XX B thuringiensis cry1Bb-related expression cassette protein SeqID12.
 DE insecticidal protein; plant; pesticide; gene therapy;
 XX lepidopteran insect pest; transgenic plant;
 KW insect infestation resistance; monocot; dicot; cry1Bb.
 KW
 XX Bacillus thuringiensis.
 OS Synthetic.
 OS
 XX WO2004020636-A1.
 PN
 XX 11-MAR-2004.
 PD
 XX 26-AUG-2003; 2003WO-US026510.
 PF
 XX 29-AUG-2002; 2002US-0407428P.
 PR
 XX (MONS) MONSANTO TECHNOLOGY LLC.
 PA (BOGD/) BOGDANOVA N N.
 PA (ROMA/) ROMANO C P.
 XX
 PI Bogdanova NN, Romano CP;
 XX WPI; 2004-269221/25.
 DR N-PSDB; ADK98488.
 DR
 XX New polynucleotide sequence optimized for expression of an insecticidal
 PT protein in a plant, useful in the control of lepidoptera insect pests,
 PT and for producing transgenic plants with the ability to resist insect
 PT infestations.
 PT
 XX Claim 7; SEQ ID NO 12; 138pp; English.
 PS
 XX This invention relates to a novel polynucleotide sequence optimised for
 CC expression of an insecticidal protein in a plant. The invention may be
 CC useful for the production of pesticides whilst the disclosed sequences
 CC may be used for gene therapy. The polynucleotide sequence and methods are
 CC useful in the control of lepidopteran insect pests, and for producing
 CC transgenic plants with the ability to resist insect infestations. The
 CC invention provides polynucleotide sequences with enhanced, improved and
 CC optimised expression in monocot and dicot plant species. The present
 CC sequence is that of a B thuringiensis cry1Bb expression cassette protein
 CC which is related to the invention.
 CC
 XX Sequence 1230 AA;

Query Match 80.8%; Score 5233.5; DB 8; Length 1230;
 Best Local Similarity 79.8%; Pred. No. 0;
 Matches 982; Conservative 94; Mismatches 149; Indels 5; Gaps 3;

Qy 2 TSNRKNELINALSTPAVSNHSTOMDLSDPARIEDSLCIAEGNNINPLYSASTVQTGIN 61
 Db 3 TSNRKNELINALSTPAVSNHSTOMDLSDPARIEDSLCIAEGNNINPLYSASTVQTGIN 62
 Qy 62 IAGRIILGVLPAGQIASFYSLVGLWPRGRDQWEIFLEHVQLINQOITENARNTAL 121
 Db 63 IAGRIILGVLPAGQIASFYSLVGLWPRGRDQWEIFLEHVQLINQOITENARNTAI 122
 Qy 122 ARIOGLGDSFRAYQOQSLEDWLENRRDARTSVLYTYQVIALELDFLNAMPFAIRNQEVPL 181
 Db 123 ARLEGLGRGYSYQQALETWLDNRNDARSRSIILERYVALELDITTAIPLFRIRNEEVPL 182
 Qy 182 LMVYAQAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVEQTRDYSQCYEWNTNTGLNS 241
 Db 183 LMVYAQAANLHLLLRDASLFGSEWMASSDQYQOEIRYTEESNHCQWNTNGLNN 242
 Qy 242 LRGTNAAASVRYNQFRRLDTLGLVDLVALFPSTRTYPIINTSAQLTREVVTDAIGATGV 301
 Db 243 LRGTNAAESWLYNQFRRLDTLGLVDLVALFPSTRTYPIINTSAQLTREIYTDPIGRNA 302
 Qy 302 --NMASMNWYNNNAPSASIAETAVIRSPHLLDFLEQLTIFSTSRWSATRHMTYVRGHTI 359
 Db 303 PSGFASNTWNNNAPSASIAEAAIFRPPHLLDPPEQLTIYSASSRWSSTQHMYWVGHRL 362
 Qy 360 QSRPIGGGLNTSTHGST-NTSINPVRLSFPSSRDVYWTESYAGVLLWGIYLEPIHGVTVR 418
 Db 363 NFRPIGGTLNTSTQGLTNTSINPVTQLQFTSRDVRTTESNAGTNI--LFTTPVNGPWAR 420
 Qy 419 FNFRNPQNTFERGTANYSQPYBSPGLQDKDSELPETTERPNYESYSHRLSHIGLISQ 478
 Db 421 FNFINPQNTYERGATYSQPYQGVGIQDFSELPETTERPNYESYSHRLSHIGLIIIG 480
 Qy 479 SRVHPVYVWTHRSADRTNTISSDITQIPLVKSFNLSGTSVVGPGFTGGDIIRTNVN 538
 Db 481 NTLRAPVYVWTHRSADRTNTIGPNRITQIPLVXALNLHSGVTVVGPGFTGGDILRRNT 540
 Qy 539 GSVLSMGLNFNTSLQRYRVRYAASQTMVLRVTGSGSTTFDQGFSTTSANESLSQS 598
 Db 541 GTFGDILRLNINPLSQRYRVRYASTDQLQFTTRINGTTVNTGNFSTWNRGNLBYRS 600
 Qy 599 FRFAEPVVGISASGSGOTAGISISNAGROTFFHPKIEFIPITATFEAEYDLERAQAVNA 658
 Db 601 FRTAGSTPFLNAQSTFTLGAQSPNSQEVYIDRVFVPAEVTFEAEYDLERAQAVNA 660
 Qy 659 LFTNTPRRLKTDVTDVHIDQVSNLVACLSDFCLEKRELLKVKYAKLSDBRNLLQD 718
 Db 661 LFTSTNPRRLKTDVTDVHIDQVSNMVACLSDFCLEKRELLFEKVYAKLSDBRNLLQD 720
 Qy 719 PNFTSINKQPDPISTNEQSNFTSIHQSEHGWSGSENIITQEGNDVFKENYVTLPGTFNE 778
 Db 721 PNFTFISGQLSPASIDQGSNFPSSINELSEHGWSGANVTIQEGNDVFKENYVTLPGTFNE 780
 Qy 779 CYPYLYQKIGESSELKAYTRYQLRGYIEDSQDLEIYLIRYNAXHETLDVPGTESWPLSV 838
 Db 781 CYPNLYQKIGESSELKAYTRYQLRGYIEDSQDLEIYLIRYNAXHETLDVPGTSLWPLSV 840
 Qy 839 ESPIGRCGEPRNCAPHEWNPDLDCSCRDGKCAHSHHFSLDIDIGCTDLHENLGVWV 898
 Db 841 ESPIGRCGEPRNCAPHEWNPDLDCSCRDGKCAHSHHFTLDDVCTDLHENLGVWV 900
 Qy 899 FKIQTQEGHARLGNLFIEBKPLLGEALSRVKAQKWRDKREKJQLETXRVYTEAKEAV 958
 Db 901 FKIQTQEGYARLGNLFIEBKPLIIGEALSRVKAQKWRDKREKJQLETXRVYTEAKEAV 960
 Qy 959 DALFVDSQYRLQADNTNIGMHAADKLVRHRIEAYLSELSVIPGVNAEIPFEELEGRILTA 1018
 Db 961 DALFVDSQYDQLQADNTNIGMHAADKLVRHRIEAYLSELVPIGVNAEIPFEELEGRILTA 1020
 Qy 1019 ISLYDARNVVKNGDFNGLTCWNVKGHDVQQSHRSVLVIPWEAEVSQAVRVCPRGY 1078
 Db 1021 MSLYDARNVVKNGDFNGLTCWNVKGHDVQQSHRSVLVIPWEAEVSQAVRVCPRGY 1080
 Qy 1079 ILRVTAKEGEGECVTIHEIENNTDELKFKNCBEEVYPTDCTCNDYTAHQGTAVCNS 1138

Db 1081 ILRVTAKEGEGCVTHIEIENNTDELKFNCEEEVYPTDTGTCNDYTAHQGTACNS 1140
QY 1139 RNAGYEDAYEVDTTASVNYKPTYEEETTDYRRDNHCEYDRGYVNYPLPAGYVTKLEY 1198
Db 1141 RNAGYEDAYEVDTTASVNYKPTYEEETTDYRRDNHCEYDRGYVNYPPVAGYVTKLEY 1200
QY 1199 FPETDKWIEIGETEGKFIVDSVELLLMEE 1228
Db 1201 FPETDTWIEIGETEGKFIVDSVELLLMEE 1230

Search completed: April 21, 2005, 04:12:16
Job time : 133 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2005, 00:32:39 ; Search time 134 Seconds
(without alignments)
4692.785 Million cell updates/sec

Title: US-10-614-524-2
Perfect score: 6479
Sequence: 1 LTSNRKNEIINALSIPAV.....IGTEGKFIVDSVELLMEE 1228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5924.5	91.4	1227	1 CIBE_BACTU	O85805 bacillus th
2	5745	88.7	1228	1 CIBA_BACTK	P05517 bacillus th
3	5740	88.6	1228	2 O93775	O93775 bacillus th
4	5737	88.5	1228	2 O93NM5	O93NM5 bacillus th
5	5235.5	80.8	1229	1 C1BB_BACTU	O45739 bacillus th
6	5214.5	80.5	1231	1 C1BD_BACTZ	O93az25 bacillus th
7	5188.5	80.1	1231	2 Q8KN72	Q8KN72 bacillus th
8	4959.5	76.5	1233	1 C1BC_BACTM	O45774 bacillus th
9	4284.5	66.1	1215	1 C1KA_BACTM	O45715 bacillus th
10	3977.5	61.4	1174	1 C1EB_BACTA	O03745 bacillus th
11	3901.5	60.2	849	2 O6PVW8	O6PVW8 bacillus th
12	3791	58.5	1176	1 C1CB_BACTG	P56953 bacillus th
13	3579	55.2	1170	1 C1JB_BACTU	O45716 bacillus th
14	3556	54.9	1167	1 C1JA_BACTU	O45738 bacillus th
15	3495.5	54.0	1189	1 C1CA_BACTE	O5518 bacillus th
16	3495.5	54.0	1189	2 O6YNE8	O6YNE8 bacillus th
17	3485.5	53.8	1189	2 Q9LSV8	Q9LSV8 bacillus th
18	3455.5	53.3	1180	2 Q9SSV8	Q9SSV8 bacillus th
19	3433.5	53.0	1165	1 C1DA_BACTA	P19415 bacillus th
20	3433	53.0	1166	1 C1GA_BACTU	O45746 bacillus th
21	3428	52.9	1179	1 C1AD_BACTA	Q03744 bacillus th
22	3426.5	52.9	1181	1 C1AE_BACTL	O03748 bacillus th
23	3426	52.9	1171	2 O06894	O06894 bacillus th
24	3426	52.9	1171	2 Q71TW6	Q71TW6 bacillus th
25	3425	52.9	1177	2 O6EIX3	O6EIX3 bacillus th
26	3421	52.8	1171	1 C1EA_BACTX	O57458 bacillus th
27	3413.5	52.7	1174	2 O45749	O45749 bacillus th
28	3413.5	52.7	1176	2 Q7WZT9	Q7WZT9 bacillus th
29	3412.5	52.7	1174	1 C1FA_BACTA	O03746 bacillus th
30	3411	52.6	1160	1 C1DB_BACTU	O45747 bacillus th
31	3405.5	52.6	1176	2 Q45736	O45736 bacillus th

32	3401	52.5	1160	2 Q93TF9	Q93TF9 bacillus th
33	3400.5	52.5	1176	1 C1AA_BACTK	P02965 bacillus th
34	3399.5	52.5	1178	1 C1AC_BACTK	P05068 bacillus th
35	3399.5	52.5	1178	2 O6XLN7	O6XLN7 bacillus th
36	3398	52.4	1177	2 O8GLY5	O8GLY5 bacillus th
37	3393.5	52.4	1178	2 O45768	O45768 bacillus th
38	3392	52.4	1177	2 O03743	O03743 bacillus th
39	3391.5	52.3	1178	2 Q9R826	Q9R826 bacillus th
40	3388.5	52.3	1176	2 Q9RC30	Q9RC30 bacillus th
41	3387	52.3	1177	2 O45735	O45735 bacillus th
42	3384	52.2	1169	1 C1FB_BACTM	O66377 bacillus th
43	3364	51.9	1177	2 O71JF1	O71JF1 bacillus th
44	3356	51.8	1169	2 O8GHE8	O8GHE8 bacillus th
45	3355	51.8	1169	1 C1GB_BACTZ	Q93az26 bacillus th

ALIGNMENTS

RESULT 1
CIBE_BACTU ID CIBE_BACTU STANDARD; PRT; 1227 AA.
AC O85805;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Pesticidal crystal protein cryIbE (Insecticidal delta-endotoxin
DE CryIB(e)) (Crystalline entomocidal protoxin) (139 kDa crystal protein).
GN Name-cryIbE; Synonyms=158C2B, cryIB(e);
OS Bacillus thuringiensis.
OC Plasmid pMYC2383.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-18872 / PS158C2;
RA Payne J.M., Cummings D.A., Cannon R.J.C., Narva K.E., Stelman S.;
RT "Bacillus thuringiensis genes encoding lepidopteran-active toxins."
RL Patent number US5723758, 03-MAR-1998.
CC -!- FUNCTION: Promotes colloidotomic lysis by binding to the midgut
epithelial cells of many lepidopteran larvae.
CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
sporulation and is accumulated both as an inclusion and as part of
the spore coat.
CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
terminus.
CC -!- SIMILARITY: Belongs to the delta endotoxin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF07326; AAC32850.1; -
CC HSSP; P07130; 1DLC.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_like.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_N; 1.
KW Plasmid; Sporulation; Toxin.
SQ SEQUENCE 1227 AA; 139084 MW; CBA847BEA0B34CD3 CRC64;

Query Match 91.4%; Score 5924.5; DB 1; Length 1227;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 1128; Conservative 37; Mismatches 60; Indels 5; Gaps 3;

QY 1 LTSNRKNEIINALSIPAVNSHTQMDLSFDARIEDSLCIAEGNINPLVSTVQTGI 60

```

Db      1  MTSNRKNEBIIALSIPAVSNHSAQNLSTDARIEDSLCIAEGNNIDPFVSASTVGTGI 60
Qy      61  NIAGRIILGVLPVAGQIAGFYSLVGLWELWPRGRDQWEI FLEHVEQLINQOI TENARNTA 120
Db      61  NIAGRIILGVLPVAGQIAGFYSLVGLWELWPRGRDQWEI FLEHVEQLIRQQVTEENTRDTA 120
Qy      121  LARLOGLGDSFRAYQOQSLDLEWLNDRDARTSRVLTQYVIALELDPLNAMPFAIRNOEVP 180
Db      121  LARLOGLGDSFRAYQOQSLDLEWLNDRDARTSRVLTQYVIALELDPLNAMPFAIRNOEVP 180
Qy      181  LLMYAQAANLHLLLRDASLFGSEFGLTSGOEIQRYYERQVEQTRDYSDYCVWEYNTGLN 240
Db      181  LLMYAQAANLHLLLRDASLFGSEFGLTSGOEIQRYYERQVEQTRDYSDYCVWEYNTGLN 240
Qy      241  SLRGTHAASVRVYNQFRDITLGLVDLVALFPSPYDTRTYPINTSAQLTRVYTDATGATG 300
Db      241  NLRGTHAASVRVYNQFRDITLGLVDLVALFPSPYDTRTYPINTSAQLTRVYTDATGATG 300
Qy      301  V--NMASMNWYNNAPSFSALETAVIRSPHLLDPLEOLTIFPSTSSRSATRHTMYNRGHT 358
Db      301  APSGFATNFWNNAPSFSALEAARVIRPHLLDPLEOLTIFPSTSSRSATRHTMYNRGHT 358
Qy      359  IQSRPIGGGLNTSHGTSNTSINPVLRSFFSRDYWTESYAGVLLWGIYLEPIHGVTVR 418
Db      359  IQSRPIGGGLNTSHGTSNTSINPVLRSFFSRDYWTESYAGVLLWGIYLEPIHGVTVR 418
Qy      418  LESRTIRGSLSTSHGTSNTSINPVLRSFFSRDYWTESYAGVLLWGIYLEPIHGVTVR 418
Db      418  LESRTIRGSLSTSHGTSNTSINPVLRSFFSRDYWTESYAGVLLWGIYLEPIHGVTVR 418
Qy      419  FNFRNPONTERTGANTYQYSPGLQKDSSETLPETTERNYESYHRLSHIGLISQ 478
Db      419  FNFRNPONTERTGANTYQYSPGLQKDSSETLPETTERNYESYHRLSHIGLISQ 478
Qy      478  FNFRNPONTERTGANTYQYSPGLQKDSSETLPETTERNYESYHRLSHIGLISQ 478
Db      478  FNFRNPONTERTGANTYQYSPGLQKDSSETLPETTERNYESYHRLSHIGLISQ 478
Qy      479  SRVHPVYVSWTHRSADRTNISDSITQIPLVKSFNLSGTSVVGPGFTGGDIIRTNVN 538
Db      479  SRVHPVYVSWTHRSADRTNISDSITQIPLVKSFNLSGTSVVGPGFTGGDIIRTNVN 538
Qy      538  GSVLSMGLNFNTSLQRYRVRYAASQTMVLRVTVGGSTTFDQGPSTMSANESLTSQS 598
Db      538  GSVLSMGLNFNTSLQRYRVRYAASQTMVLRVTVGGSTTFDQGPSTMSANESLTSQS 598
Qy      598  FRAEPFVGISASGQTAGISINAGROTHFDFKIEFIPITATFAEYDLERAQAVNA 658
Db      598  FRAEPFVGISASGQTAGISINAGROTHFDFKIEFIPITATFAEYDLERAQAVNA 658
Qy      658  LFTNTPRRLKTVTDVHIDQVSNLVAFLDSEFLDEKRELLKVKYAKRLSDERNLLQD 718
Db      658  LFTNTPRRLKTVTDVHIDQVSNLVAFLDSEFLDEKRELLKVKYAKRLSDERNLLQD 718
Qy      718  PNFTSINKQPDFISTNRQSNFTSIHQSEHGWSGSENIITQEGNDVFKENYVILPGTFNE 778
Db      718  PNFTSINKQPDFISTNRQSNFTSIHQSEHGWSGSENIITQEGNDVFKENYVILPGTFNE 778
Qy      778  CYPTLYQKIGSELKAYTRYQRYGIEDSDQDLEIYLIRNAKHETLDVPGTESVWPLSV 838
Db      778  CYPTLYQKIGSELKAYTRYQRYGIEDSDQDLEIYLIRNAKHETLDVPGTESVWPLSV 838
Qy      838  ESPIGRGCEPNRCAPFENWPDLDCCRDCEKCAHSHHPSFLDIDIGCTDLHENLGVWV 898
Db      838  ESPIGRGCEPNRCAPFENWPDLDCCRDCEKCAHSHHPSFLDIDIGCTDLHENLGVWV 898
Qy      898  FKIKTQEGHARLGNLFEIEBKPLLGALSRVKAERKWRDKREKLETKRVTYEAKEAV 958
Db      898  FKIKTQEGHARLGNLFEIEBKPLLGALSRVKAERKWRDKREKLETKRVTYEAKEAV 958
Qy      958  DALFVDSQYRDLQADTNIHMAADKLVRIRAYLSLSVIPGVNAEIPFEELEGRITTA 1018
Db      958  DALFVDSQYRDLQADTNIHMAADKLVRIRAYLSLSVIPGVNAEIPFEELEGRITTA 1018
Qy      1018  ISLYDARNVVKNDENNGLACWNVKGVHDVQOQSHHRSVLVIPWEAEVSAQVRCVGRGY 1077
Db      1018  ISLYDARNVVKNDENNGLACWNVKGVHDVQOQSHHRSVLVIPWEAEVSAQVRCVGRGY 1077
Qy      1078  ILRVAYKEGYGCGCVTIHEIENNTDELKFKNCEEEVYPTDGTGNDYTAHQGTAVCNS 1138

```

```

Db      1078  ILRVAYKEGYGCGCVTIHEIENNTDELKFKNCEEEVYPTDGTGNDYTAHQGTAVCNS 1137
Qy      1138  RNAGYEDAEVDVTASVNYKPYTEETTYDVRDNDHCEYDRGVNYPPLPAGWTMKELEY 1198
Db      1138  RNAGYEDAEVDVTASVNYKPYTEETTYDVRDNDHCEYDRGVNYPPLPAGWTMKELEY 1197
Qy      1198  FPETDKVWIEIGTEGKFIIVDSVELLMEE 1228
Db      1198  FPETDKVWIEIGTEGKFIIVDSVELLMEE 1227

RESULT 2
C1BA_BACTK STANDARD; PRT; 1228 AA.
ID C1BA_BACTK STANDARD; PRT; 1228 AA.
AC P05517; Q45731;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Pesticidal crystal protein cryIIa (Insecticidal delta-endotoxin
DE CryIIa(a)) (Crystalline entomocidal protoxin) (140 kDa crystal protein).
DE Names-cryIIa; Synonyms-cryA4, cryIIa(a);
GN Bacillus thuringiensis (subsp. kurstaki), and
OS Bacillus thuringiensis (subsp. entomocidus).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=29339, 1436;
RN [1]
SEQUENCE FROM N.A.
RC SPECIES=B.t.kurstaki; STRAIN=HD-2;
RA MEDLINE=88203216; PubMed=3362680;
RX Brizzard B.L., Whiteley H.R.;
RT "Nucleotide sequence of an additional crystal protein gene cloned from
RT Bacillus thuringiensis subsp. thuringiensis.";
RL Nucleic Acids Res. 16:2723-2723 (1988).
RN [2]
SEQUENCE FROM N.A.
RC SPECIES=B.t.entomocidus; STRAIN=HD-110;
RA Soetaert P.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Promotes colloidomotic lysis by binding to the midgut
CC epithelial cells of insects.
CC -1- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -1- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -1- SIMILARITY: Belongs to the delta endotoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X06711; CAA29898.1; -
CC EMBL; X95704; CAA65003.1; -
CC FIR; S00873; S00873.
CC HSP; P07130; IDLC.
CC InterPro; IPR001178; Endotoxin.
CC InterPro; IPR005638; endotoxin_C.
CC InterPro; IPR005639; endotoxin_N.
CC InterPro; IPR008979; Gal_bind_like.
CC Pfam; PF03944; Endotoxin_C; 1.
CC Pfam; PF00555; Endotoxin_M; 1.
CC Pfam; PF03945; Endotoxin_N; 1.
CC Sporulation; Toxin.
FT VARIANT 150 150 Y -> H (in strain HD-110).
SQ SEQUENCE 1228 AA; 139647 MW; C8E3A19B5D98575 CRC64;

Query Match 88.7%; Score 5745; DB 1; Length 1228;
Best Local Similarity 89.1%; Pred. No. 0;
Matches 1098; Conservative 35; Mismatches 90; Indels 10; Gaps 3;

```

```
Qy 1 LSNRKNENEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MTSNRKNENEIIN-----AVSNHSAQMDLLDPDARIEDSLCIAEGNNIDPPFVSASTVQTGI 55
Qy 61 NIAGRIILGVLPAGPAGIASFYSFLVGLWPRGRDQWEIPLFHEVQELINQOITENARNTA 120
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
56 NIAGRIILGVLPAGPAGIASFYSFLVGLWPRGRDQWEIPLFHEVQELINQOITENARNTA 115
Qy 121 LARIQLGLSDSFAYQQSLEDWLENRDARTSRVLYTQYIALELDFLAMPFLAIRNQSEVP 180
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 LARIQLGLSDSFAYQQSLEDWLENRDARTSRVLYTQYIALELDFLAMPFLAIRNQSEVP 175
Qy 181 LLMYAQAANLHLLLDASLFGSEFGLTSQEIORYYERQVETRDYSDYCVWYNTGLN 240
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
176 LLMYAQAANLHLLLDASLFGSEFGLTSQEIORYYERQVETRDYSDYCVWYNTGLN 235
Qy 241 SLRGTAASWVRVYNQFRDRLTLGLVDLVALPSPDYTRTPINTSAQLTREVTDAIGATG 300
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
236 SLRGTAASWVRVYNQFRDRLTLGLVDLVALPSPDYTRTPINTSAQLTREVTDAIGATG 295
Qy 301 VNMAWMNNYNNAPSFSIAETAVIRSPHLLDFLEQLTIFSTSSRSWSTRHMTYWRGHTIQ 360
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
296 VNMAWMNNYNNAPSFSIAETAVIRSPHLLDFLEQLTIFSTSSRSWSTRHMTYWRGHTIQ 355
Qy 361 SRPIGGGLNTSTHGSTNTSINPVLRSFSDRVYVWTSYAGVLLWGIYLEPIHGVTYFRFN 420
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
356 SRPIGGGLNTSTHGATNTSINPVLRFASRDVYRTSYAGVLLWGIYLEPIHGVTYFRFN 415
Qy 421 FRNPQTFERTANYSPYSPGLQKDSFELPPTERTERNYESYSHRSLHIGLISQSR 480
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
416 FTNPQNTSDRGATANYSPYSPGLQKDSFELPPTERTERNYESYSHRSLHIGLISQSR 475
Qy 481 VHVPIVSWTHRSADRTNTISDSITQPLVKSFMNLSGTSVVSQPGFTGGDIIRTNVGS 540
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
476 VNVPIVSWTHRSADRTNTIGNRNTQIPWVKASLPQGTTVVRGPGFTGGDILARTWGG 535
Qy 541 VLSMGLFNNTSLQRYRVRYAASQTVLRVTYVGGSTTFDQGPSTMSANESLTSQSR 600
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
536 FGPIRVTVNGPLTKRYRIGFYASTVDFDFVSRGGTIVANFRFLRTWNSGDELKYNFV 595
Qy 601 FAEPFVGISAGSQ-TAGISISNAGQTFHFDKIEPIPTATFEASYDLERAQEVNAL 659
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
596 RRAEFTTFTFTQIQDIIRTSIQGLSGNGEVYIDKIEIIPVATFEASYDLERAQEVNAL 655
Qy 660 FTNTNPRRLKTDVTDYHIDQVSNLVACLSDFCLEKRELLEKVKYAKRLSDERNLLQDP 719
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
656 FTNTNPRRLKTDVTDYHIDQVSNLVACLSDFCLEKRELLEKVKYAKRLSDERNLLQDP 715
Qy 720 NPTSINKQPDFISTNEQSNFTSIHQSEHGWWGSENITIQBGNDVFKENYVTLPGTFNEC 779
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
716 NPTSINKQPDFISTNEQSNFTSIHQSEHGWWGSENITIQBGNDVFKENYVTLPGTFNEC 775
Qy 780 YPTLYYOKIGESLKVATRYQLRGYIEDSQDLEIYLYRYNAKHETLDVPGTESWPLSVE 839
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
776 YPTLYYOKIGESLKVATRYQLRGYIEDSQDLEIYLYRYNAKHETLDVPGTESWPLSVE 835
Qy 840 SPICRCCEPNRCAPHFENWPDLCSCRDGEKCAHSHHFLSDIDIGCTDLHENLGVWVF 899
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
836 SPICRCCEPNRCAPHFENWPDLCSCRDGEKCAHSHHFLSDIDIGCTDLHENLGVWVF 895
Qy 900 KIKTOEGHARLGNLEFIEEKPPLGALSVRVRAEKWRDKREKLQLETKRYVTEAKBAVD 959
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
896 KIKTOEGHARLGNLEFIEEKPPLGALSVRVRAEKWRDKREKLQLETKRYVTEAKBAVD 955
Qy 960 ALFVDSQVNRLOADNTGMIHAADKLVRHREAYLSELVTPGVNAEIFEELGRITAI 1019
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
956 ALFVDSQVNRLOADNTGMIHAADKLVRHREAYLSELVTPGVNAEIFEELGRITAI 1015
Qy 1020 SLYDARNVKNQDFNGLACNNVKGHVDVQOSHRSVLVPEWEAEYSQAVRVCPCGGYI 1079
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1016 SLYDARNVKNQDFNGLTCNVNKGHVDVQOSHRSVLVPEWEAEYSQAVRVCPCGGYI 1075
```

```
Qy 1080 LRVYAYKEGEGCVTTHIEINNTDELKFNCEEEVYPTDGTGTCNDYTAHQGTA----V 1135
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1076 LRVYAYKEGEGCVTTHIEINNTDELKFNCEEEVYPTDGTGTCNDYTAHQGTAHQGTA 1135
Qy 1136 CNSNAGYDAYEYDVTASVNYKPYTBEETVTVRRDNHCEYDRGVNYPPLPAGYMTKE 1195
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1136 CNSNAGYDAYEYDVTASVNYKPYTBEETVTVRRDNHCEYDRGVNYPPLPAGYMTKE 1195
Qy 1196 LEYPPETDKWIEIGETEGKPIVDSEVLLMEE 1228
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1196 LEYPPETDKWIEIGETEGKPIVDSEVLLMEE 1228
Qy 1228 LEYPPETDKWIEIGETEGKPIVDSEVLLMEE 1228
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 3
Q93T75 PRELIMINARY; PRT; 1228 AA.
ID Q93T75 AC Q93T75
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Delta-endotoxin CryIa2.
GN Name=CryIa2;
OS Bacillus thuringiensis (subsp. entomocidus).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBITaxID=1436;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HD-9;
RA Mat Isa M.N., Abdullah M.A.F., Mahadi N.M.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF363025; AAK51084.1; -.
DR HSP; P07130; 1DLC.
DR GO; GO:0005102; F:receptor binding; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001178; Endotoxin_C.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_like.
DR Pfam; PF03944; Endotoxin_C_1.
DR Pfam; PF00555; Endotoxin_M_1.
DR Pfam; PF03945; Endotoxin_N_1.
SQ SEQUENCE 1228 AA; 139620 MW; 3DA2A4DBF59C95C3 CRC64;
Query Match 88.6%; Score 5740; DB 2; Length 1228;
Best Local Similarity 89.0%; Pred. No. 0;
Matches 1097; Conservative 36; Mismatches 90; Indels 10; Gaps 3;
Qy 1 LSNRKNENEIINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MTSNRKNENEIIN-----AVSNHSAQMDLLDPDARIEDSLCIAEGNNIDPPFVSASTVQTGI 55
Qy 61 NIAGRIILGVLPAGPAGIASFYSFLVGLWPRGRDQWEIPLFHEVQELINQOITENARNTA 120
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
56 NIAGRIILGVLPAGPAGIASFYSFLVGLWPRGRDQWEIPLFHEVQELINQOITENARNTA 115
Qy 121 LARIQLGLSDSFAYQQSLEDWLENRDARTSRVLYTQYIALELDFLAMPFLAIRNQSEVP 180
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 LARIQLGLSDSFAYQQSLEDWLENRDARTSRVLYTQYIALELDFLAMPFLAIRNQSEVP 175
Qy 181 LLMYAQAANLHLLLDASLFGSEFGLTSQEIORYYERQVETRDYSDYCVWYNTGLN 240
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
176 LLMYAQAANLHLLLDASLFGSEFGLTSQEIORYYERQVETRDYSDYCVWYNTGLN 235
Qy 241 SLRGTAASWVRVYNQFRDRLTLGLVDLVALPSPDYTRTPINTSAQLTREVTDAIGATG 300
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
236 SLRGTAASWVRVYNQFRDRLTLGLVDLVALPSPDYTRTPINTSAQLTREVTDAIGATG 295
Qy 301 VNMAWMNNYNNAPSFSIAETAVIRSPHLLDFLEQLTIFSTSSRSWSTRHMTYWRGHTIQ 360
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
296 VNMAWMNNYNNAPSFSIAETAVIRSPHLLDFLEQLTIFSTSSRSWSTRHMTYWRGHTIQ 355
Qy 361 SRPIGGGLNTSTHGSTNTSINPVLRSFSDRVYVWTSYAGVLLWGIYLEPIHGVTYFRFN 420
```


Db 356 SRPIGGGLNTSTHGTATNTSINPVTLRFASRDVYRTESYAGVLLAGVILEPIHGVPTVRFN 415
Qy 421 FRNPQNTFERGTANYSQPYESPGQLKDSFETLPPETTERNYESYSHRLSHIGLIQSQR 480
Db 416 FTNPQNTSDRGRTANYSQPYESPGQLKDSFETLPPETTERNYESYSHRLSHIGLIQSQR 475
Qy 481 VHVVPYSWTHRSADRTNTISDSITQIPLVKSFNLSGTSVVGPGFTGGDIIRTNVGS 540
Db 476 VNVVPYSWTHRSADRTNTIGNRIITQIPWKASLPGQTTVVRGPGFTGGDIIRTNVGS 535
Qy 541 VLSMGLNFNTSLQRYRVRYAASQTMVLRVTVGGSTTFDQGPSTMSANESLTSQSF 600
Db 536 FGPIRVTVNGPLTQRYRIGRYASTVDFDFVSRGGTTVNNFRFLRTMNSGDELKYGNFV 595
Qy 601 FAEPFVGISASGQ-TAGISISNNAQGTFFHDKIEPIPTATFEAYDYDLERAQEAVAL 659
Db 596 RRAFTTPTFTQIQTIRTSIQGLSGNGEVYDKIEIIPVTATFEAYDYDLERAQEAVAL 655
Qy 660 FTNTNPRRLKTDVTDYHIDQVSNLVACLDEFCLDEKRELLEKVKYAKRLSDERNLLQDP 719
Db 656 FTNTNPRRLKTDVTDYHIDQVSNLVACLDEFCLDEKRELLEKVKYAKRLSDERNLLQDP 715
Qy 720 NFTSINKQPDFISTNEQSNFTSIHQSEHGWMGSENITIOEGNDVFKENYVTLPGTFNEC 779
Db 716 NFTSINKQPDFISTNEQSNFTSIHQSEHGWMGSENITIOEGNDVFKENYVTLPGTFNEC 775
Qy 780 YPTVLYKIGESLKAATRYQLRGYIEDSQLEIYLRYNKAKHETLDVPGTESWPLSV 839
Db 776 YPTVLYKIGESLKAATRYQLRGYIEDSQLEIYLRYNKAKHETLDVPGTESWPLSV 835
Qy 840 SPIGRCEPNRCAPHFENWPDLCSCRDGKCAHSHHFLDIDIGCTDLHENLGVWVVF 899
Db 836 SPIGRCEPNRCAPHFENWPDLCSCRDGKCAHSHHFLDIDIGCTDLHENLGVWVVF 895
Qy 900 KIKTQEGHARLGNLEFIEEKPGLLEALSRVKRAEKKWRDKREKLQLETKRYVTEAKEAVD 959
Db 896 KIKTQEGHARLGNLEFIEEKPGLLEALSRVKRAEKKWRDKREKLQLETKRYVTEAKEAVD 955
Qy 960 ALFVDSQYNRLQADTNGIMHAADKLVHRTREAYLSLSELPVGNABIFEELEGRHITAI 1019
Db 956 ALFVDSQYNRLQADTNGIMHAADKLVHRTREAYLSLSELPVGNABIFEELEGRHITAI 1015
Qy 1020 SLVDARNVKNGDPNNGLACVNVKGVHDVQSHRSVLVPEAEVSVQAVRCPGRGYI 1079
Db 1016 SLVDARNVKNGDPNNGLACVNVKGVHDVQSHRSVLVPEAEVSVQAVRCPGRGYI 1075
Qy 1080 LRVYAYKEGYEGCVTTHIEENNTDELKFNKBEEVYPTDGTGNDYTAHQGTAA----V 1135
Db 1076 LRVYAYKEGYEGCVTTHIEENNTDELKFNKBEEVYPTDGTGNDYTAHQGTAA----V 1135
Qy 1136 CNSRNAGYEDAYEDVTTASVNYKPYEETVTVDRDNHCEYDRGYVNYPLPAGYMTKE 1195
Db 1136 CNSRNAGYEDAYEDVTTASVNYKPYEETVTVDRDNHCEYDRGYVNYPPVAGYVTK 1195
Qy 1196 LEVFPETDKVWIEIGETEGKFIVDSEVLLMEE 1228
Db 1196 LEVFPETDKVWIEIGETEGKFIVDSEVLLMEE 1228

RESULT 4
Q93NM5 ID Q93NM5 PRELIMINARY; PRT; 1228 AA.
AC Q93NM5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cry1Ba.
GN Name=Cry1Ba;
OS Bacillus thuringiensis
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
.RN [1]

SEQUENCE FROM N.A.
RA Zhang J., Song F., Huang D.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF368257; AAK63251.1; -;
DR HSP; P07130; IDLC.
DR GO; GO:0005102; F:receptor binding; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal bind like.
DR Pfam; PF03944; Endotoxin_C; 1.
DR Pfam; PF00555; Endotoxin_M; 1.
DR Pfam; PF03945; Endotoxin_W; 1.
SQ SEQUENCE 1228 AA; 139665 MW; E86D9842341FB439 CRC64;

Query Match 88.5%; Score 5737; DB 2; Length 1228;
Best Local Similarity 89.0%; Pred. No. 0;
Matches 1097; Conservative 35; Mismatches 91; Indels 10; Gaps 3;

Qy 1 LTSNRKNENIINALSIPAVSNHSTQMDLSPPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db 1 MTSNRKNENIIN-----AVSNHSAQMDLLPDARIEDSLCIAEGNNIDPFVSASTVQTGI 55
Qy 61 NIAGRILGVLPFAGQIASFYSLVAGELWPRGDQWEIFLEHVEQLINQOITENARNTA 120
Db 56 NIAGRILGVLPFAGQIASFYSLVAGELWPRGDQWEIFLEHVEQLINQOITENARNTA 115
Qy 121 LARLOGLGDSFRAYQOSLEDWLENRDDARTSRVLYTYQYIALELDFLNAMPLFAIRNQEVP 180
Db 116 LARLOGLGDSFRAYQOSLEDWLENRDDARTSRVLYTYQYIALELDFLNAMPLFAIRNQEVP 175
Qy 181 LMVYAQAANLHLLLRDASLFGSEFGLTSQETQRYRQVETRDYSDYCVWEYNTGLN 240
Db 176 LMVYAQAANLHLLLRDASLFGSEFGLTSQETQRYRQVETRDYSDYCVWEYNTGLN 235
Qy 241 SLRGTTNAASVWRVYQFRRLDTLGLVLDLVALFPSYDTRTPINTSAQLTRVYTDALGATG 300
Db 236 SLRGTTNAASVWRVYQFRRLDTLGLVLDLVALFPSYDTRTPINTSAQLTRVYTDALGATG 295
Qy 301 VNMAWMNWNNAAPSFSAIETAVIRSPHLDFLEQLTIFSTSSRWASATRMHTYWRGTIQ 360
Db 296 VNMAWMNWNNAAPSFSAIETAVIRSPHLDFLEQLTIFSTSSRWASATRMHTYWRGTIQ 355
Qy 361 SRPIGGGLNTSTHGTATNTSINPVTLRFASRDVYRTESYAGVLLAGVILEPIHGVPTVRFN 420
Db 356 SRPIGGGLNTSTHGTATNTSINPVTLRFASRDVYRTESYAGVLLAGVILEPIHGVPTVRFN 415
Qy 421 FRNPQNTFERGTANYSQPYESPGQLKDSFETLPPETTERNYESYSHRLSHIGLIQSQR 480
Db 416 FTNPQNTSDRGRTANYSQPYESPGQLKDSFETLPPETTERNYESYSHRLSHIGLIQSQR 475
Qy 481 VHVVPYSWTHRSADRTNTISDSITQIPLVKSFNLSGTSVVGPGFTGGDIIRTNVGS 540
Db 476 VNVVPYSWTHRSADRTNTIGNRIITQIPWKASLPGQTTVVRGPGFTGGDIIRTNVGS 535
Qy 541 VLSMGLNFNTSLQRYRVRYAASQTMVLRVTVGGSTTFDQGPSTMSANESLTSQSF 600
Db 536 FGPIRVTVNGPLTQRYRIGRYASTVDFDFVSRGGTTVNNFRFLRTMNSGDELKYGNFV 595
Qy 601 FAEPFVGISASGQ-TAGISISNNAQGTFFHDKIEPIPTATFEAYDYDLERAQEAVAL 659
Db 596 RRAFTTPTFTQIQTIRTSIQGLSGNGEVYDKIEIIPVTATFEAYDYDLERAQEAVAL 655
Qy 660 FTNTNPRRLKTDVTDYHIDQVSNLVACLDEFCLDEKRELLEKVKYAKRLSDERNLLQDP 719
Db 656 FTNTNPRRLKTDVTDYHIDQVSNLVACLDEFCLDEKRELLEKVKYAKRLSDERNLLQDP 715
Qy 720 NFTSINKQPDFISTNEQSNFTSIHQSEHGWMGSENITIOEGNDVFKENYVTLPGTFNEC 779
Db 716 NFTSINKQPDFISTNEQSNFTSIHQSEHGWMGSENITIOEGNDVFKENYVTLPGTFNEC 775

```

Qy 780 YPTLYOKIGESSELKAYTRYQLRGYIEDSODLEIYLIRYNAKHETLDVPGTESVPLSVE 839
Db 776 YPTLYOKIGESSELKAYTRYQLRGYIEDSODLEIYLIRYNAKHETLDVPGTESVPLSVE 835
Qy 840 SPIGRCEPNRCAPHFENPDLDSCSDGKCAHSHHFSIDIDIGCTDLHENLGVWVVF 899
Db 836 SPIGRCEPNRCAPHFENPDLDSCSDGKCAHSHHFSIDIDIGCTDLHENLGVWVVF 895
Qy 900 KIKTOEGHARLGNLEFTEEPKLLGEALSRVRAEKWRDKREKLQLETKRVTYTAKEAVD 959
Db 896 KIKTOEGHARLGNLEFTEEPKLLGEALSRVRAEKWRDKREKLQLETKRVTYTAKEAVD 955
Qy 960 ALFVDSQYRLQADTNGIMHAADKLVHRIEAYLSVLPVGNABIFEELEGRITAI 1019
Db 956 ALFVDSQYRLQADTNGIMHAADKLVHRIEAYLSVLPVGNABIFEELEGRITAI 1015
Qy 1020 SLVDARNVKNGDNFNGGLACVWVKGHDVQOOSHRSVLVPEWEAEVSQAVRVCPCRGYI 1079
Db 1016 SLVDARNVKNGDNFNGGLACVWVKGHDVQOOSHRSVLVPEWEAEVSQAVRVCPCRGYI 1075
Qy 1080 LRVYAYKEGEGCVTHIEINNTDELKFKNCEEEVYPTDGTGNDYTAHQGTA---V 1135
Db 1076 LRVYAYKEGEGCVTHIEINNTDELKFKNCEEEVYPTDGTGNDYTAHQGTA---V 1135
Qy 1136 CNSRNAGYEDAYEDVTASVNYKPYTEETVTVRRDNHCEYDRGYNYPLPAGYMTKE 1195
Db 1136 CNSRNAGYEDAYEDVTASVNYKPYTEETVTVRRDNHCEYDRGYNYPLPAGYMTKE 1195
Qy 1196 LEYFPETDKWIEIGETEGKFIVDSVLLMEE 1228
Db 1196 LEYFPETDKWIEIGETEGKFIVDSVLLMEE 1228
Qy 1196 LEYFPETDKWIEIGETEGKFIVDSVLLMEE 1228
Db 1196 LEYFPETDKWIEIGETEGKFIVDSVLLMEE 1228

RESULT 5
C1BB_BACTU STANDARD; PRT; 1229 AA.
AC Q45739;
AT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE Pesticial crystal protein cryIbB (Insecticidal delta-endotoxin
DE CryIb(b)) (Crystalline entomocidal protoxin) (140 kDa crystal protein).
GN Names-cryIbB; Synonyms-cryET5, cryIb(b);
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-21110 / EGS847;
RA Donovan W.P., Tan Y., Jany C.S., Gonzalez J.M. Jr.;
RT "Bacillus thuringiensis cryet4 and cryet5 toxin genes and proteins
RT toxic to lepidopteran insects.";
RL Patent number US5322687, 21-JUN-1994.
CC -I- FUNCTION: Promotes colloidal osmotic lysis by binding to the midgut
CC epithelial cells of many lepidopteran larvae.
CC -I- DEVELOPMENTAL STAGE: The crystal protein is produced during
CC sporulation and is accumulated both as an inclusion and as part of
CC the spore coat.
CC -I- MISCELLANEOUS: Toxic segment of the protein is located in the N-
CC terminus.
CC -I- SIMILARITY: Belongs to the delta endotoxin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L32020; AAA22344.1; --
CC HSSP; P02965; IC1Y.
CC InterPro; IPR001178; Endotoxin.

```

```

DR InterPro; IPR005638; endotoxin_C.
DR InterPro; IPR005639; endotoxin_N.
DR InterPro; IPR008979; Gal_bind_like.
DR Pfam; PF03944; Endotoxin_C_1.
DR Pfam; PF05555; Endotoxin_M_1.
DR Pfam; PF03945; Endotoxin_N_1.
KW Sporulation; Toxin.
SQ SEQUENCE 1229 AA; 139769 MW; A4C949DB675C3269 CRC64;

Query Match      80.8%; Score 5235.5; DB 1; Length 1229;
Best Local Similarity 79.8%; Pred. No. 0;
Matches 982; Conservative 95; Mismatches 149; Indels 5; Gaps 3;

Qy 1 LTSNRKNEIEINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db 1 MTSNRKNEIEINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Qy 61 NIAGRIILGVLPAGQIASFYSLVGLWPRGRDWEIIFLEHVEQLINQOITENARNTA 120
Db 61 NIAGRIILGVLPAGQIASFYSLVGLWPRGRDWEIIFLEHVEQLINQOITENARNTA 120
Qy 121 LARLOGLGDSFRAYQQSLEDWLENRDDARTSRVLYTOYIALELDFLAMPFLAIRNOEVP 180
Db 121 IARLEGLRGYRSYQQALETWLDNRDARSIIILERYVALELDITTAIPLFRIRNEVP 180
Qy 181 LLMVYAAANLHLLLRDASLPGSEFGLTSGEIQRYERQVEQTRDYSDYCEVYNTGLN 240
Db 181 LLMVYAAANLHLLLRDASLPGSEMGASSDVNQYQEQIRYTEESNHCVQYNTGLN 240
Qy 241 SLRGTNAASVRYNORRDLTLGVLDLVALFPSYDTETYPINTSAOLTREVYDAIGATG 300
Db 241 NLRGTNAESWLRYNORRDLTLGVLDLVALFPSYDTETYPINTSAOLTREVYDAIGATG 300
Qy 301 V--NMASMNWYNNNAPSFAIETAVIRSPHLLDFLEQLTIFSTSSRSATRHMTYWRGHT 358
Db 301 APSGFASWNNFNNAFSAIEAAIFPPHLLDEPEQLTIYSASSRWSSTOHNNYVWVGR 360
Qy 359 IQSRPIGGGLNTSTHGST--NTSINPVRLSPFSRDVYMTESVAGVLLMGVILEPHIGVPTV 417
Db 361 LNFPRIGGLTNTSTQGLTNNTSINPVLTQTRDVRVYTESNAGTNI--LFTTPVGVVPA 418
Qy 418 RFNFRPQNTFERGTANYSQYESPGLOKDSLELTPETTERPNYESYSHRLSHIGLIS 477
Db 419 RFNFINPQNIYERGATTYQPYQGVGQLFDSLELTPETTERPNYESYSHRLSHIGLII 478
Qy 478 QSRVHVVPVYSWTHRSADRTNTISSDSITQIPLVKSFNLNSGTSVWSGPGFTGGDIIRTNV 537
Db 479 GNTLRAPVYVSWTHRSADRTNTIGPNRITQIPLVKALNLSGTVVVGPGFTGGDILRRTN 538
Qy 538 NGSVLSGLNFMNNTSLQRYRVRYVRYAASQTMVLRVTVVGSTTFPDQGFPSMTSANEISTSQ 597
Db 539 TGTFGDILRLINVPVLSQRYRVRYRYASTDLQPFTRINGTTVNIQGNFSRTMNRDNLRYR 598
Qy 598 SFRFAEPVVCISAGSQTAGISINNAGRTQTFHDKIEFIPITATFEAYDLEAQAQAVN 657
Db 599 SFTAGFSTPFNFNAQSTFTLGAQSFNSQEVYDRVFPVPAEYVTFEAYDLEAQAQAVN 658
Qy 658 ALFTNTNPRLKTDTVTDYHIDQVSNLVACLSDFCLDEKELLEKVYAKRLSDERNLLQ 717
Db 659 ALFTNTNPRLKTDTVTDYHIDQVSNLVACLSDFCLDEKELLEKVYAKRLSDERNLLQ 718
Qy 718 DPNFTSINKQPDFISTNEQSNFTSIHQSEHGWSGSENITIQEGNDVFKENYVTLPGTFN 777
Db 719 DPNFTSINKQPDFISTNEQSNFTSIHQSEHGWSGSENITIQEGNDVFKENYVTLPGTFN 778
Qy 778 ECTPYLYQKIGESSELKAYTRYQLRGYIEDSODLEIYLIRYNAKHETLDVPGTESVPLS 837
Db 779 ECTPYLYQKIGESSELKAYTRYQLRGYIEDSODLEIYLIRYNAKHETLDVPGTESVPLS 838
Qy 838 VESPIGRCEPNRCAPHFENPDLDSCSDGKCAHSHHFSIDIDIGCTDLHENLGVWV 897
Db 839 VESPIGRCEPNRCAPHFENPDLDSCSDGKCAHSHHFSIDIDIGCTDLHENLGVWV 898

```

QY 898 VFLLKTOGHARLGNLEFIEKPLIGALSRVKAERKWRDKREKLOLETKRVVTEAKEA 957
 Db 899 VFLLKTOGHARLGNLEFIEKPLIGALSRVKAERKWRDKREKLOLETKRVVTEAKEA 958
 QY 958 VDALLFVDSQYNRLQADTNIGMIHAADKLVRHREAYLSLSVIFGVNAEIPFEELEGRIT 1017
 Db 959 VDALLFVDSQYNRLQADTNIGMIHAADKLVRHREAYLSLSVIFGVNAEIPFEELEGRIT 1018
 QY 1018 AISLYDARNVKNQDFNGLACWNVKGVVDVQOSSHRSVLVLPWEAEVQAVRVCPRG 1077
 Db 1019 AISLYDARNVKNQDFNGLACWNVKGVVDVQOSSHRSVLVLPWEAEVQAVRVCPRG 1078
 QY 1078 YILRVATYKEGYGCGVTHIEINNTDELKPKNCEEEVPTDGTCTNDYTAHOGTAVCN 1137
 Db 1079 YILRVATYKEGYGCGVTHIEINNTDELKPKNCEEEVPTDGTCTNDYTAHOGTAVCN 1138
 QY 1138 SRNAGYEDAVEVDVTASVNYKPTVEEYTDVRRDNHCEYDRGVYVPPVAGVYTKLE 1197
 Db 1139 SRNAGYEDAVEVDVTASVNYKPTVEEYTDVRRDNHCEYDRGVYVPPVAGVYTKLE 1198
 QY 1198 YFPETDKVWIEIGETGEKFIVDSVELLMEE 1228
 Db 1199 YFPETDKVWIEIGETGEKFIVDSVELLMEE 1229

RESULT 6

C1BD_BACTZ STANDARD; PRT: 1231 AA.
 ID C1BD_BACTZ
 AC Q9AZS;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Peptidic crystal protein cryIAB (Insecticidal delta-endotoxin
 DE CryIAB(d)) (Crystalline entomocidal protoxin) (140 kDa crystal protein).
 GN Name-cryIAB; Synonyms-cryEI, cryIAB(d), cryIAl;
 OS Bacillus thuringiensis (subsp. whanensis).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=52024;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HD-525;
 RX MEDLINE=2015386; PubMed=10688690;
 RA Kuo W.-S., Lin J.-H., Tzeng C.-C., Kao S.-S., Chak K.-F.;
 RT "Cloning of two new cry genes from Bacillus thuringiensis subsp.
 RL whanensis strain";
 RL Curr. Microbiol. 40:227-232 (2000).
 CC -1- FUNCTION: Promotes colloid osmotic lysis by binding to the midgut
 CC epithelial cells of lepidopteran larvae. Toxic to Plutella
 CC xylostella.
 CC -1- DEVELOPMENTAL STAGE: The crystal protein is produced during
 CC sporulation and is accumulated both as an inclusion and as part of
 CC the spore coat.
 CC -1- MISCELLANEOUS: Toxic segment of the protein is located in the N-
 CC terminus.
 CC -1- SIMILARITY: Belongs to the delta endotoxin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U70726; AAD10292.1; --
 DR HSSP; P02965; IC1Y.
 DR InterPro; IPR001178; Endotoxin.
 DR InterPro; IPR005638; Endotoxin C.
 DR InterPro; IPR005639; Endotoxin N.
 DR InterPro; IPR008979; Gal bind like.
 DR Pfam; PF03944; Endotoxin_C_1.
 DR Pfam; PF00555; Endotoxin_M; 1.
 DR Pfam; PF03945; Endotoxin_N; 1.

KW Sporulation; Toxin.
 SQ SEQUENCE 1231 AA; 139654 MW; 129A0371CDDBE52 CRC64;
 Query Match 80.5%; Score 5214.5; DB 1; Length 1231;
 Best Local Similarity 79.7%; Pred. No. 0;
 Matches 987; Conservative 94; Mismatches 140; Indels 17; Gaps 9;
 QY 1 L TSNRKNENIINALSIPAVSNHSTOMDLSFDARIEDSLCIAEGNNINPLVSASTVGTGI 60
 Db 1 MTSNRKNENIINALSIPAVSNHSAQMDLSLDARIEDSLCIAEGNNINPLVSASTVGTGI 60
 QY 61 NIAGRILGVGVFPAGQASFYFLGELWPRGRDWEIEFLEHVEQLINQIITENARNTA 120
 Db 61 NIAGRILGVGVFPAGQASFYFLGELWPRGRDWEIEFLEHVEQLIRQOVVNTENTNTA 120
 QY 121 LARLOGLGDSFRAYQOOSLEDWLENRRDARTSRVLYTQYIALELDFLNAMFLFAIRNOEVP 180
 Db 121 IARLEGLRGYRSYQQALETWLDNRNDARSISILERYVALELDITTAIPLFRINEEVP 180
 QY 181 LLMVYQAANLHLLILDASLFGSEFGLTSOETQRYRQVQVOTRDYSDYCVWYNTGLN 240
 Db 181 LLMVYQAANLHLLILDASLFGSEWGMASSDVNYQYQEQIRYTEESYNSHCVQWYNTGLN 240
 QY 241 SLRGITNAASWVRYNQFRDLTLGLVDLVALFPSYDTRTYPINTSAQLTREVIYTDPIGTN 300
 Db 241 NLRGITNAESWLRYNQFRDLTLGLVDLVALFPSYDTRTYPINTSAQLTREVIYTDPIGTN 300
 QY 301 V--NMASMNWNNNAPSFSAIETAVIRSHPLDLFLQLTIFSTSSRWASRTHMYWRGHT 358
 Db 301 APSGFASNTWNNNAPSFSAIEAAIIRPRLHLLDFPEQLITYSASSRWSSTQHNNYVVGHR 360
 QY 359 IQSRPIGGGLNTSTHGST-NTSINPVRLSFFSDVWVTSYAGVLLWGLYLEIHGVPTV 417
 Db 361 LNFPRPGGLTNTSTQGLTNTSINPVTLQFTSDVYRTESNAGTNI--LFTTPVNGVPMWA 418
 QY 418 RFNFRNPQNTFERGTANYSQPYESPGQLQKDSSTELPPTETTERPNYESHRLSHLGLIS 477
 Db 419 RFNFINPQNIYERGATTSYQYGVGVLQFDSSETLPPTETTERPNYESHRLSHLGLII 478
 QY 478 QSRVHVPTVSWTHRSADRTNTISSDSITQIPLVKSFNLSNGTSVVGPGFTGGDIIRTN- 536
 Db 479 GNTLRAPVYVSWTHRSADRTNTIGPNRITQIPAVKGRFLFNG-SVIGPGFTGGDVVLNR 537
 QY 537 VNGSVLSMG-----LNFNNTSLQRVHRVRYAASQTMVLRVTVGGSTTFDQGPSTMSAN 591
 Db 538 NNGNINQRVIEVPIQFTSTIS-TRYRVRVRYASVTSIELNVNLGNSIITNTLPTAASL 596
 QY 592 ESLSQSFRFAFPVGI-SASGSQTAGISISNAGRQTFHDFDKIEFIPITATPEAEYDLE 650
 Db 597 DNLQSGDFGVVEINNAFTSATGNIVGARNFSANA---EVLIDRFEPITATPEAEYDLE 653
 QY 651 RQAEAVNALFTWNPRLKTDVTDHIDQVSNLVACLSDFECLDEKRELLKVKYAKRLS 710
 Db 654 RAQKAVNALFTSTNPRRLKTDVTDHIDQVSNVACLSDFECLDEKRELLFEKVYAKRLS 713
 QY 711 DERNLLODNFTSINKOPDFISITNEOSNFTSIHQSEHGWSGSENITIOEGNDVFKENVY 770
 Db 714 DERNLLODNFTFISQLSFASIDGQSNFTSINELSEHGWSGSENITIOEGNDVFKENVY 773
 QY 771 TLPGTNECPTYLYKIGESLKYATRYOLRGYIEDSQDLRIYLIRYNAKHETLDPVGT 830
 Db 774 TLPGTNECYPNLYKIGESLKYATRYOLRGYIEDSQDLRIYLIRYNAKHETLDPVGT 833
 QY 831 ESWPVLVSFPIGRCEPNRCAPHFENWPDLCSDGCEKCAHSHHFDLIDIGCTDLH 890
 Db 834 DSLWPLSVKSPIGRCGEPNRCAPHFENWPDLCSDGCEKCAHSHHFDLIDIVGCTDLH 893
 QY 891 ENLGVWVFEKIKTQEGHARLGNLEFIEKPLIGALSRVKAERKWRDKREKLOLETKRV 950
 Db 894 ENLGVWVFEKIKTQEGHARLGNLEFIEKPLIGALSRVKAERKWRDKREKLOLETKRV 953
 QY 951 YTEAKEAVDALFVDSQYNRLQADTNIGMIHAADKLVRHREAYLSLSVIFGVNAEIPFEE 1010

Db 954 YTEAKETVDALFVDSHYNRLQADTNIGMHAADRLVHRIHEAYLPPELPFGINAVIFEE 1013

Qy 1011 LEGRIITAIISLYDARNVVKNGDFNGLACWNVKGHDVQSHHRSVLVLPWEAEVSQAV 1070

Db 1014 LENRISTAFSLYDARNVVKNGDFNGLSCWNVKGHDVQSHHRSDLVLPWEAEVSQAV 1073

Qy 1071 RVCPRGYILLRVATYKGYGEGCVTHIEINNTDELAFKNCBEEVYPTDGTGNDYTAH 1130

Db 1074 RVCPRGYILLRVATYKGYGEGCVTHIEINNTDELAFKNCBEEVYPTDGTGNDYTAH 1133

Qy 1131 QGTAVCNRSNAGYEDAYEVDVTTASVNYKPYEEBETDVRDNDHCEYDRGVVYPPPLPAG 1190

Db 1134 QGTAAACNSRNAGYEDAYEVDVTTASVNYKPYEEBETDVRDNDHCEYDRGVVYPPVAG 1193

Qy 1191 YMTKELEYFPETDKWIEIGETGKFIIVDSVELLMEE 1228

Db 1194 YVTKELEYFPETDVTWIEIGETGKFIIVDSVELLMEE 1231

RESULT 7

Q8KNY2 PRELIMINARY; PRT; 1231 AA.

AC Q8KNY2; 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE CryIBII.

GN Name=CryIBII;

OS Bacillus thuringiensis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1428;

RN [1]

RP SEQUENCE FROM N.A.

RA Iakova I.A., Isaakov Y.B., Rymar' S.E., Yarovoi S.V.;

RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.

EMBL: AY138457; AA093496.1;

DR HSP; P02965; iCIY.

DR GO: 0005102; F:receptor binding; IEA.

DR GO: 0006952; P:defense response; IEA.

DR GO: 0009405; P:pathogenesis; IEA.

DR InterPro: IPR001178; Endotoxin.

DR InterPro: IPR005638; endotoxin_C.

DR InterPro: IPR005639; endotoxin_N.

DR InterPro: IPR008979; Gal bind like.

DR Pfam: PF03944; Endotoxin_C; 1.

DR Pfam: PF00555; Endotoxin_M; 1.

DR Pfam: PF03945; Endotoxin_N; 1.

SQ SEQUENCE 1231 AA; 139764 MW; C9F2848A9297EA00 CRC64;

Query Match 80.1%; Score 5188.5; DB 2; Length 1231;

Best Local Similarity 79.3%; Pred. No. 2.4e-317;

Matches 982; Conservative 97; Mismatches 142; Indels 17; Gaps 9;

Qy 1 LTSNRKNEIINALSTIPAVSNHSTQMDLSPDARIEDSLCIAEGNNPLVSASTVGTGI 60

Db 1 MTSNRKNEIINALSTIPVSNPSTQNLSPDARIEDSLCAEVANNIDPFVSASTVGTGI 60

Qy 61 NIAGRIILGVLPVGPAGQIAGFYSLVGLMWRGRDOWEIEFLHVEQLINQOITENARNTA 120

Db 61 NIAGRIILGVLPVGPAGQIAGFYSLVGLMWRGRDOWEIEFLHVEQLINQOITENARNTA 120

Qy 121 LARLQGLGDSFRAYQOSLEDWLENRDARTSRVLYTQYIALELDPLNAMPLFAIRNOEVP 180

Db 121 IARLEGLGRVRSYQQALETWLDNRNDARSIIILERYVALELDITTAIPLFIRIRNEEVP 180

Qy 181 LMVYAQAANLHLLLRDASLFGSEFGLTQSETQRYRQYERQETRDYSVCVWNTGLN 240

Db 181 LMVYAQAANLHLLLRDASLFGSEWGWASSNQYQEQIRYTEESNHCVQWNTGLN 240

Qy 241 SLRGNTAASVRVNOFRDRLTLGVLDLVALPPSYDTEPTINTSAQLTREYTDALGATG 300

Db 241 NLRGNTAASWLRYNQFRDRLTLGVLDLVALPPSYDTEPTINTSAQLTREYTDPIGRTN 300

Qy 301 V--NMASMWYNNAPSFAIETAVIRSPHLLDLEOLTIFSTSSRWSATRHMTYWRGHT 358

Db 301 APGFASTFNFNAPSFAIETAVIRSPHLLDLEOLTIFSTSSRWSATRHMTYWRGHT 360

Qy 359 IQSRPIGGGLNTSTHGST-NTSINPVRLSFFSRDVMYTESYAGVLLMGITYLEPIHGVPV 417

Db 361 LNFPRPGTGLNTSTOGLTNTSINPVTLOFTSRDVMYTESYAGVLLMGITYLEPIHGVPV 418

Qy 418 RFNFRPNQNTFERGTANYSPYSPGLQKDSFETELPETERPNYSESYSYHRLSHIGLIS 477

Db 419 RFNFRPNQNTFERGTANYSPYSPGLQKDSFETELPETERPNYSESYSYHRLSHIGLIS 478

Qy 478 QSRVHVVPVSWTHRSADRTNTISSDSTOJPLVKSFNLSGTSVVSQGGTGGDIIRTN- 536

Db 479 GNTLRAPVYSWTHRSADRTNTIGNRTIQIPAVKGRFLFNG-SVISGPGFTGGGVRLNR 537

Qy 537 VNGSVLSMG----LNFNNTSLQRYRVRVRYAASQTMVLRVTVGGSTTFDQGFPTMSAN 591

Db 538 NNGNIQNRGYIEVPIQFTSTS-TRYRVRVYASVTSIELNVNLSNSSFINTLTPATAASL 596

Qy 592 ESITSQSFRFAEPFVGI-SASGSQTAGISINNAGROTFFHDKIEFIPITATPEAEVDLE 650

Db 597 DNLQSGDFGYVEINNAFTSATGNIVGARNFSA---EVIIDRPEFIPVATPEAEVDLE 653

Qy 651 RAQEAUNALFTNPRRLKTDVTDYHIDQVSNLVACLSDFCLDKRELLEKVKYAKRLS 710

Db 654 RAQKAVNALFTSNPRRLKTDVTDYHIDQVSNLVACLSDFCLDKRELLEKVKYAKRLS 713

Qy 711 DERNLQDPNFTSINKOPDFISTNEQSNFTSIHQSHGWMGSENITIQGNDVFKENYV 770

Db 714 DERNLQDPNFTSINKOPDFISTNEQSNFTSIHQSHGWMGSENITIQGNDVFKENYV 773

Qy 771 TLPCTFNECYTYLYQKIGSELKAYTRYQLRGYIEDSQDLEIYLIRYNAKHETLDVPGT 830

Db 774 TLPCTFNECYTYLYQKIGSELKAYTRYQLRGYIEDSQDLEIYLIRYNAKHETLDVPGT 833

Qy 831 ESWPLSVESPIGRCEPNRCAPHFENWPDLCSCRDGECACAHSHHFSIDIDGCTDLH 890

Db 834 DSLWPLSVKSPIGRCCEPNRCAPHFENWPDLCSCRDGECACAHSHHFSIDIDGCTDLH 893

Qy 891 ENLGVWVVKIKTOEGHARLGNLEFTEKPLIGALSRRVRAEKWRDKREKLQLETKRV 950

Db 894 ENLGVWVVKIKTOEGHARLGNLEFTEKPLIGALSRRVRAEKWRDKREKLQLETKRV 953

Qy 951 YTEAKEAVDALFVDSQYNRLQADTNIGMHAADRLVHRIHEAYLPPELPFGINAVIFEE 1010

Db 954 YTEAKEAVDALFVDSHYNRLQADTNIGMHAADRLVHRIHEAYLPPELPFGINAVIFEE 1013

Qy 1011 LEGRIITAIISLYDARNVVKNGDFNGLACWNVKGHDVQSHHRSVLVLPWEAEVSQAV 1070

Db 1014 LENRISTAFSLYDARNVVKNGDFNGLSCWNVKGHDVQSHHRSDLVLPWEAEVSQAV 1073

Qy 1071 RVCPRGYILLRVATYKGYGEGCVTHIEINNTDELAFKNCBEEVYPTDGTGNDYTAH 1130

Db 1074 RVCPRGYILLRVATYKGYGEGCVTHIEINNTDELAFKNCBEEVYPTDGTGNDYTAH 1133

Qy 1131 QGTAVCNRSNAGYEDAYEVDVTTASVNYKPYEEBETDVRDNDHCEYDRGVVYPPPLPAG 1190

Db 1134 QGTAAACNSRNAGYEDAYEVDVTTASVNYKPYEEBETDVRDNDHCEYDRGVVYPPVAG 1193

Qy 1191 YMTKELEYFPETDKWIEIGETGKFIIVDSVELLMEE 1228

Db 1194 YVTKELEYFPETDVTWIEIGETGKFIIVDSVELLMEE 1231

RESULT 8

C1BC BACTM

ID C1BC BACTM STANDARD; PRT; 1233 AA.

AC Q45774;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Pesticidal crystal protein cryIbC (Insecticidal delta-endotoxin)

QY 433 ANYSQVSPGLQKDSLETLPPTERRPNVSYSHRLSHGLISQSRVHV-----PVSWS 488
 DB 406 V-----DSLTELPPEDNSVPPREGYSHRLCHATFVQSGTLPPLTTGPFSW 451
 QY 489 THRSADRTNTISSDSITQIPLVKFNLNSGTSVVGSGFTGCDIIRTNVNGSVLSMGLNF 548
 DB 452 THRSATDRNIIPVQINQIPLVKAFNITSGTSVVGSGFTGCDIIRTNVNGSVLSMGLNF 511
 QY 549 NNTSLQVRVRVRAASQTMVLRVTGSGTTFDQGFPTSMANESLTSQSRFPAEPFVGI 608
 DB 512 SNTTLQVRVRVRAASQTMVLRVTGSGTTFDQGFPTSMANESLTSQSRFPAEPFVGI 571
 QY 609 SASGSQTAGISINNAGRTQPHFKIEFIPITATFEAYDLERAQEAVALFTNTNPRRL 668
 DB 572 SASGSQASISINNVGRQMFHLDRIEFLPVTSTFEEDYDLERAQEAVALFTSTNQLGL 631
 QY 669 KTDVTDHIDVSNLVACLSEFCLEKRELEKLVKAKRLSDRNLLOPNFTSINKQP 728
 DB 632 KTDVTDHIDVSNLVACLSEFCLEKRELEKLVKAKRLSDRNLLOPNFTSINKQP 691
 QY 729 DFISTNQSNFTSITHEQSEHGWSGSENIITQEGNDVFKENVVTLPGTFNFCYPTLYOKI 788
 DB 692 D-----RGRWGSTDTITQGGDDVFKENVVTLPGTFNFCYPTLYOKI 733
 QY 789 GESELKAYTRYQLRGYIEDSQDLIELYIRYNKAKHETLDVPGTESVWPLSVESPIGRGCEP 848
 DB 734 DESKLSYTRYELRGYIEDSQDLIELYIRYNKAKHEIVNVPCTGSLWPLSIENSIGPCGEP 793
 QY 849 NRCAPHEWNPDLDCSRDGEKCAHSHHSFLSDIDIGCTDLHENLGVVWVFKIKTQGHGA 908
 DB 794 NRCAPHEWNPDLDCSRDGEKCAHSHHSFLSDIDIGCTDLHENLGVVWVFKIKTQGHGA 853
 QY 909 RLGNLEFIEKPLLGEALSRVKRAEKWKDRKLEKQLETKRVVYTEAKEAVDALFVDSQYN 968
 DB 854 RLGNLEFIEKPLLGEALSRVKRAEKWKDRKLEKQLETKRVVYTEAKEAVDALFVDSQYN 913
 QY 969 RLQADTNIGMIHAADKLVHRIREAYLSLSVPGVNAEIEFEELEGRITIAISLYDARNV 1028
 DB 914 RLQADTNIGMIHAADKLVHRIREAYLSLSVPGVNAEIEFEELEGRITIAISLYDARNV 973
 QY 1029 KNGDPNGLACVWVKGVDV-QQSHHSVLVI PWEEAEVSQAVRVCGRGYILRVAYKE 1087
 DB 974 KNGDPNGLACVWVKGVDV-QQSHHSVLVI PWEEAEVSQAVRVCGRGYILRVAYKE 1033
 QY 1088 GYGGCVTIEIENNTDELKFKNCEEEVYPTDTGTCNDYTAHQ- --GTAVCSNRNAGYE 1144
 DB 1034 GYGGCVTIEIENNTDELKFKNCEEEVYPTDTGTCNDYTAHQ- --GTAVCSNRNAGYE 1092
 QY 1145 DAYEVDTTASVNYKPYTEETTYTDVRRDNHCEYDRGVYVNPPLPAGYMTKELEYFPETDK 1204
 DB 1093 KSYESNSVQADYASVYEEKADTDGRRDNHCSNRNGYGDVYTPLPAGYVTKLEYFPETDK 1152
 QY 1205 VWIEIGTEGKFIYDVSVELLIMEE 1228
 DB 1153 VWIEIGTEGKFIYDVSVELLIMEE 1176

RESULT 13

C1JB_BACTU STANDARD; PRT; 1170 AA.
 AC Q45716;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Pesticidal crystal protein cryIJB (Insecticidal delta-endotoxin
 CryIJB(b)) (crystalline entomocidal protoxin) (134 kDa crystal protein).
 GN Name=cryIJB; Synonyms=cryET1, cryIJB(b);
 OS Bacillus thuringiensis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1428;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-18973 / EG5092;

RA von Tersch M.A., Gonzalez J.M. Jr.;
 RT "Bacillus thuringiensis cryET1 toxin gene and protein toxic to
 RL lepidopteran insects";
 RL Patent number US5356623, 18-OCT-1994.
 CC -!- FUNCTION: Promotes colloidotomotic lysis by binding to the midgut
 CC epithelial cells of many lepidopteran larvae. Toxic to Plutella
 CC xylostella.
 CC -!- DEVELOPMENTAL STAGE: The crystal protein is produced during
 CC sporulation and is accumulated both as an inclusion and as part of
 CC the spore coat.
 CC -!- MISCELLANEOUS: Toxic segment of the protein is located in the N-
 CC terminus.
 CC -!- SIMILARITY: Belongs to the delta endotoxin family.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U31527; AAA98959.1; -;
 CC HSPSP; P02965; 1C1Y.
 DR InterPro; IPR0011178; Endotoxin.
 DR InterPro; IPR005638; endotoxin_C.
 DR InterPro; IPR005639; endotoxin_N.
 DR InterPro; IPR008979; Gal_bind_like.
 DR Pfam; PF03944; Endotoxin_C; 1.
 DR Pfam; PF00555; Endotoxin_M; 1.
 DR Pfam; PF03945; Endotoxin_N; 1.
 KW Sporulation; Toxin.
 SQ SEQUENCE 1170 AA; 133553 MW; B053B9619B78DC19 CRC64;
 Query Match 55.2%; Score 3579; DB 1; Length 1170;
 Best Local Similarity 60.5%; Pred. No. 4.1e-216;
 Matches 725; Conservative 125; Mismatches 275; Indels 73; Gaps 17;
 QY 46 NINPL-VSASTVGTGINIAGRIILGVLPFAGQIASFYSFLVGLWPRGDOWEIFLEHV 104
 DB 31 DIDPLEVSNLLOFLNNF-----VPGGGFISGLFKIWCALRP---SDNELFLAQI 79
 QY 105 EQLINQOITENARNTALARLQGLDSFRAYQQSLEDWLENDRDARTSRVLYTQYIALEL 164
 DB 80 EQLIDQRIEATVRAKIAELEGLGRSQLYVEAFKEWEETPDNTAARSRTVERFRIDAQ 139
 QY 165 FLNAMPFLAIRNOEVLMLVYQAQANLHLLLDASLFGSEFGLTQEQIQRVYERQEQT 224
 DB 140 IEANIFSPRIQPGFEVFLSVYQAQANLHLLLDASLFGSEFGLTQEQIQRVYERQEQT 199
 QY 225 RYSDYCVQWYNTGLNLRGTNAASWVRVYNOFRDLTLGLVLDLVALFPSTYDRTYPINTS 284
 DB 200 HEYSDHCVTYKTELELGTSTRAQWKIYNQFRRLTLVLDIVAVFPNVGDKLYPIQTX 259
 QY 285 AQLTREYVTDALGATGVNMAWMNNYNNAPSAIET-AVIRSPHLLDLFLEQLTIFSTSS 343
 DB 260 SELTREIYTPV--SEVYGAINNQN-----GIQTERQIRQPHLMDPENTMTVTSYN 312
 QY 344 R-----WSATRHMTYWGHTIQSRPIGGGLNTSTGNTSINPVLSPFSDRVVWTESYA 399
 DB 313 RREYVWSGLEMTAYFTGFA-----GPQVSFPFAGTRGDAAPPFNVRVNDGVI----- 360
 QY 400 GVLLWGIYLEPIHGVTFRFNRPNQPTFERGTANYSQP-----YESFGLQKDSLETLP 455
 DB 361 RIUSAPFYAPPLG--TSVLGSRGEFMF--ALNNISPPPSARYRNP--TVDSLVSIPP 414
 QY 456 ETTERPNYESYSHRLSHGLISQSRVHVYVSWTHRSADRTNTISSDSITQIPLVKSNFL 515
 DB 415 QDNSVPPHRRGSSHRLSHVWTRNSS----PIFHWTHRSATTTNTRINSVDITQIPMWKAYNL 470
 QY 516 NSGTSVVGSGFTGGDIIRTNVNGSVLSMGLNFNTSLQRYVRVRYAASQTMVLRVTVG 575
 DB 471 HAGATVVRGPGFTGGDILRRTSGTMVTLKVDASAVRNQRIYRFRYATSNFYFVVRRG 530

Db 160 YAQANLHLA1LRDSVIFGERWGLTTINNVENYNRLIRHIDEYADHCANTYNRGLNNLPK 219
Qy 245 TNAASWRYNQFRDLTLGVLDLVALPSPYDTRYPINTSAQLTREYVYTAIGATGVNMA 304
Db 220 STYQDMITYNRLRDLTLVLDIAAFPFPYDNRYPYIQPVGQLTREYVYDPL-----INFN 275
Qy 305 SMWYNNNAPSFAIETAVIRSHLLDFLEOLTIFSTSSRWSATRHMTYMRGHTIQSRPI 364
Db 276 PQLOSVALPFTFNWESSAIRNPHLFDILNNLTIFTD---WFSVGRNPFYMGGRHVISSLI 332
Qy 365 GGGGLTSTHGSTNTSINPRLSPFSRDVYWTESYAGVLL-----WGIYLEPIHGVPTVRFN 420
Db 333 GGGNITSPIYCREANQPPRSFTENGVPFRTLNPTRLRLQQPWPAPPFNLRGVGVGEFS 392
Qy 421 FRNPQTFE---RGTANYQBPYESPGLQDKDSELTPEPETERPNYESYSHRLSHIGLIS 477
Db 393 --TPTNSFTYRGCTV-----DSLTELPEONSVPPEGYSYHRLCHATFVQ 436
Qy 478 QSRVHV---PVYSWTHRSADRTWTISSDSITQIPLVKSNNLSNGTSVVSQPGFTGGDII 533
Db 437 RSGTPEFTTGVVSWTHRSATLTNTIDPERINQIPLVKGFRVWGTSVITGPGFTGGDIL 496
Qy 534 RTNVNGSVLSMGLNFNNTSLQRYVRVRYAASQ-----TMVLRTVYGGSTTFDQGFPS 586
Db 497 RRNTFGDFSVLOVNNINSPIQRYELRFRYASSRDARVIVLTGAASTGVGGQSVNMPLOK 556
Qy 587 TMSANESLTSOSFPABFP-----VGISASGSQTAGISISNNAGRQTFHFDKIE 635
Db 557 TMEIGENLTSRTFRYTDSPNPFSPRANPDIIGISEQPLFGAG-SISG-----ELYIDKIE 611
Qy 636 FIPITATFEAYDLERAQEAVALFTWNPRLRATDVTYHIDQVSNLVACLSDDEFCLDE 695
Db 612 IILADATFEASDRLERAQKAVNALFTSSNQIGLKTDTYDHIQVSNLVCLSDDEFCLDE 671
Qy 696 KRELLEKVKYAKRLSDERNLQDPNFTSINKOPDFISTNEQSNFTSHEQSEHGHWGSEN 755
Db 672 KRELSEKVKAKRLSDERNLQDPNFGINRQPD-----RCWRGSTD 713
Qy 756 ITIQEGNDVPKENYVTLPGTFNECYTYLYOKIGSELKAYTRYQLRGYIEDSDLEIYL 815
Db 714 ITIQGGDDVPKENYVTLPGTVDECYTYLYOKIDESKLKAYTRYELRGYIEDSDLEIYL 773
Qy 816 IRYNAKHETLDVPGETESVWPLSVESPIGRGCEPNRCAPHFENWPDLDSCRCDEKCAHHS 875
Db 774 IRYNAKHEIVNVPGTGLWPLSAQSPICKGCEPNRCAPHLEWNPDLDCSCRDGKCAHHS 833
Qy 876 HPSLDIDIGCTDLHENLGVWVFKIKTOEGHARLGNLEPTEBKPLLGEALSRVKRAEKK 935
Db 834 HHFTLIDVGCTDLNEDLGWVIFKIKTQGHARLGNLEFLEEKPLLGEALARVKRAEKK 893
Qy 936 WRDKREKLOLETKRVTYTEAKEAVDALFVDSQYNRLOADTNIGMIHAADKLVRHIREAYLS 995
Db 894 WRDKREKLOLETNIVYKEAKESVDALFVNSQYDRLOQVDTNIAIMIHAADKRVHIREAYLP 953
Qy 996 ELSVIPGVNAIEIPELEGRIITAILSLYDARNVKNQDFNNGLACWNVKGHVDV--QOSHHR 1054
Db 954 ELSVIPGVNAAIPELEGRIIFTAYSLYDARNVIXNGDFNNGLLCWNVKGHVDVEEQNNHR 1013
Qy 1055 SVLVIPWEAEVSOAVRVCGRGYILRVAYKEGYGEGCVTHIEINNTDELKFNCEEE 1114
Db 1014 SVLVIPWEAEVSEVRVCGRGYILRVAYKEGYGEGCVTHIEIEDNTDELKFSNCVEE 1073
Qy 1115 EYVPTDTGTCDNYTA---HQTAVCNRNAGYDAYEDVDTASVNYKPTYEEETYTDDR 1170
Db 1074 EYVFNNTVTCNNYTGTOEYEGT--YTSRNGYDEAYGNPNPSVPADYASVYEEKSYTDGR 1131
Qy 1171 RDNHCEYDRGVNTPPLPAGWMTKELEYFPBPTDKWIEIGETEGKFTVDSVELLLMEE 1228
Db 1132 RENPCESNRGYGDTPLPAGVYTKDLEYPFETDKWIEIGETEGFTVDSVELLLMEE 1189

Search completed: April 21, 2005, 04:16:15

Job time : 139 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2005, 16:15:00 ; Search time 15262 Seconds
(without alignments)
11705.830 Million cell updates/sec

Title: US-10-614-524-1

Perfect score: 3687
Sequence: 1 ttgactcaaataggaaaaa.....tactccttaggaggaatag 3687

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3687	100.0	3687	6	AX189649
2	3666.2	99.4	4173	6	AR359382 Sequence
3	3666.2	99.4	4173	6	AX098667 Sequence
4	3266.6	88.6	3684	1	AF077326 Bacillus
5	3266.6	88.6	3684	6	I90319 Sequence 7
6	3260.2	88.4	3684	6	AR359400 Sequence
7	3260.2	88.4	3684	6	AX098692 Sequence
8	3136.8	85.1	3687	1	X06711 B. thuringi
9	3135.2	85.0	3687	1	AF368257 Bacillus
10	3133.6	85.0	4074	1	BTCRV1BA2
11	3133.6	85.0	4074	6	I42103 Sequence 1
12	3133.6	85.0	4798	1	AF363025 Bacillus
13	2687.2	72.9	3934	1	BACCRYIE
14	2687.2	72.9	3934	6	AR068828 Sequence
15	2687.2	72.9	3934	6	I38760 Sequence 3
16	2687.2	72.9	3934	6	I70138 Sequence 3
17	2630.2	71.3	3627	6	AX138580 Sequence
18	2630.2	71.3	3627	6	AX146772 Sequence
19	2579.2	70.0	8306	1	BTU07026 Bacillus th

20	2563.2	69.5	3696	1	AY138457
21	2544.4	69.0	3663	6	AX088012 Sequence
22	2541.2	68.9	3663	6	AX088010 Sequence
23	2384.2	64.7	3624	6	AX138578 Sequence
24	2384.2	64.7	3624	6	AX146770 Sequence
25	2271.6	61.6	3842	1	BTMCRYIBC
26	2145	58.2	2549	1	AY570734 Bacillus
27	2004.8	54.4	3651	6	AX138576 Sequence
28	2004.8	54.4	3651	6	AX146768 Sequence
29	2000.4	54.3	4623	1	BTU28801
30	1796.4	48.7	3525	1	BACCRYIEB
31	1793.4	48.6	3522	6	AR085762 Sequence
32	1793.4	48.6	3522	6	I11609 Sequence 4
33	1793.4	48.6	3522	6	I73040 Sequence 1
34	1791.8	48.6	3522	6	AR362261 Sequence
35	1790.2	48.6	3522	6	A27529 B.thuringie
36	1662	45.1	4106	1	BACCRYIC
37	1662	45.1	4106	6	I82521 Sequence 1
38	1662	45.1	4106	6	I93759 Sequence 1
39	1581.6	42.9	3558	6	AR451656 Sequence
40	1581.6	42.9	3558	6	BD090512 Plant opt
41	1501	40.7	3766	6	BTU31527
42	1495.6	40.6	3624	6	AR029509 Sequence
43	1495.6	40.6	3624	6	AR098462 Sequence
44	1495.6	40.6	3624	6	BD194455 Chimeric
45	1495.6	40.6	3624	6	I41422 Sequence 6

ALIGNMENTS

RESULT 1
AX189649
LOCUS AX189649 3687 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from Patent WO0147952.
ACCESSION AX189649
VERSION AX189649.1 GI:15143038
KEYWORDS
SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1
AUTHORS Arnaut, G., Boets, A., Damme, N., Mathieu, E., Vanneste, S. and van Rie, J.
TITLE Insecticidal proteins from Bacillus thuringiensis
JOURNAL Patent: WO 0147952-A 1 05-JUL-2001;
Aventis CropScience N.V. (BE)
FEATURES
source Location/Qualifiers
1..3687
/organism="Bacillus thuringiensis"
/mol_type="unassigned DNA"
/db_xref="taxon:1428"
1..3687
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC50778.1"
/db_xref="GI:15143039"
/translation="MTSNRKNENEINALSIPAVSNHSTQWDLSPDARIEDSLCIAEC
NNINPLVSTVGTGINAGRIVLGVGFAGQIASFYSLVGLWPRGRDQWEIFLE
HVEQLNQITENARTALRLOGLQDSFRAYQQSLDLEWLRDADRTRISVLYTQYIA
LELDFLNMPFLAIRNQEVLPLVYQAQAHLLILLDASLFGSEFGTSQEIQRYVE
ROYEOTRDYSDYCVENYNTGLNSLRGTNAASWRYNQFRDLTGVLDLVALPSPYDT
RTYPTNSAQLTREVTDAITGATGVNWSMNWYNNAPSASAJETAVIRSPHLLDLE
QLTFTSSRSWASRTMYRHTIQSRPIGGGLNTSTHGSTNTSINPRLSFRSDV
WYTESVAGVLLGWIIYLEIPHTVFRNFENPQNTFERGTANSPQYSPESGLQKDE
TELPPTETTERPNYESHRLSHCLISQSRVHVSVSWTHRSADRTNTISDSITQIP
LVKSFNLNCGTSVVSQGFQGGDIIRTNVNGSVLSMGLNFNTSLQRYVRVRAASQ
TWLVRTVVGSTTFDQGFPMSTMSANESLTSQSFRAEFPGVIGISASQTAGISNNA
GROTFPHDKIEFIPITATFAEYDLERAQAVNALFTNTPRRLKTDVTDYHIDQVSN
LVACLSEDFCLDEKLEKLVKAKLSRDLNLLQDPNFTSINKQPFISTNEQSNFT
SIHQEHEHGWGSGENITIQEGNDVFKENYVTLPGTFNECYPTLYLQKIGESELKAYTR

YQLRGYIEDSODLEIYLIRYNAKHETLDPGTSVPLSVESPIGRCEPNRCAPHE
 WNPDCSCRDGKCAHSHHSLDIDIGCTDLHENLGVWVFKIKTOEGHARLGNLE
 FIEBKPLGLBALSVKAEKWRDKREKLOLETKRVYTEAKEAVDALFVDSOYNRLQA
 DTNIGMHAADKLVHRIREAYLSLSVPGWNAEIPFELSGRIITALSILYDARNVKN
 GDFNGLACWNKGVHDVQOQSHRSLVLIPEWEABVSQAVRPGRYILIRVAYKEG
 YGBCVTHIEINNTDELKFNCEEEVEYPTDGTGNDYTAHQGTAVCNBSRNYEDA
 YEVDTTASVNYKPTYPEETVTVDRNDHCEYDRGVNYPPLPAPGMYMKELEYPPEDK
 VWIEIGTEGKFIVDSVELLLMBE"

ORIGIN

		Query Match	100.0%;	Score 3687;	DB 6;	Length 3687;		
		Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;	Gaps 0;	
		Matches 3687;	Conservative 0;					
Qy	1	TTGACTTCAAAATAGAAAATGAGAAATGAAATATAAATGCTTTATCGATTCCAGCTGTA	60					
Db	1	TTGACTTCAAAATAGAAAATGAGAAATGAAATATAAATGCTTTATCGATTCCAGCTGTA	60					
Qy	61	TGGAATCAITCCACACAAATGGATCTATCACAGATGCTGTAATGAGGATTTCTTTGT	120					
Db	61	TGGAATCAITCCACACAAATGGATCTATCACAGATGCTGTAATGAGGATTTCTTTGT	120					
Qy	121	ATAGCCAGGGGAATAATATCAATCCACTTGTAGCCCATCAACAGTCCAAAACGGGTATT	180					
Db	121	ATAGCCAGGGGAATAATATCAATCCACTTGTAGCCCATCAACAGTCCAAAACGGGTATT	180					
Qy	181	AACATAGCTGTAGAACTACTAGGTGTATTAGGGTACCGTTGCTGCACAAATAGCTAGT	240					
Db	181	AACATAGCTGTAGAACTACTAGGTGTATTAGGGTACCGTTGCTGCACAAATAGCTAGT	240					
Qy	241	TTTTATAGTTTTCTTGTGTGTAATTATGCCCOCGCGCAGAGATCAGTGGGAAATTTTC	300					
Db	241	TTTTATAGTTTTCTTGTGTGTAATTATGCCCOCGCGCAGAGATCAGTGGGAAATTTTC	300					
Qy	301	CTAGAACATGTGCAACAACTTATAAATCAACAAATAACAGAAAATGTAGGAAATACGGCA	360					
Db	301	CTAGAACATGTGCAACAACTTATAAATCAACAAATAACAGAAAATGTAGGAAATACGGCA	360					
Qy	361	CTTGCTCGATTACAAGTTTAGGAGATTCCTTTAGAGCTTATCAACAGTCACTTGAAGAT	420					
Db	361	CTTGCTCGATTACAAGTTTAGGAGATTCCTTTAGAGCTTATCAACAGTCACTTGAAGAT	420					
Qy	421	TGGCTAGAAAACCGTGATGATGCAAGAAACGAGAAAGTGTCTTTATACCCAAATATAGCC	480					
Db	421	TGGCTAGAAAACCGTGATGATGCAAGAAACGAGAAAGTGTCTTTATACCCAAATATAGCC	480					
Qy	481	TTAGAACTTGATTTCTTAAATGCGATGCGCTTTTCGAAATAGAAACCAAGAAATTCOA	540					
Db	481	TTAGAACTTGATTTCTTAAATGCGATGCGCTTTTCGAAATAGAAACCAAGAAATTCOA	540					
Qy	541	TTTAAATGATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGCTCT	600					
Db	541	TTTAAATGATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGCTCT	600					
Qy	601	CTTTTGGTATGAAATTTGGGCTTTACATCGCAGGAAATTCACGTTATTATGAGCGCAA	660					
Db	601	CTTTTGGTATGAAATTTGGGCTTTACATCGCAGGAAATTCACGTTATTATGAGCGCAA	660					
Qy	661	GTGGACAAACGAGAGATTATTCGACTATTGCTAGAAATGGTATATACAGGTCTAAAT	720					
Db	661	GTGGACAAACGAGAGATTATTCGACTATTGCTAGAAATGGTATATACAGGTCTAAAT	720					
Qy	721	AGCTTGAGGGGCAAAATGCGCAAGTTGGGTGCTTATAATCAATTCCTGAGAGATCTA	780					
Db	721	AGCTTGAGGGGCAAAATGCGCAAGTTGGGTGCTTATAATCAATTCCTGAGAGATCTA	780					
Qy	781	ACGTTAGGGGATTTAGATCTAGTGGCACTATTCCAAAGCTATGACACTCCGCTTATCCA	840					
Db	781	ACGTTAGGGGATTTAGATCTAGTGGCACTATTCCAAAGCTATGACACTCCGCTTATCCA	840					
Qy	841	ATAAATACGAGTCTCAGTTTAAACAGGGAAGTTTATACAGACCGAATTTGAGCAACAGG	900					
Db	841	ATAAATACGAGTCTCAGTTTAAACAGGGAAGTTTATACAGACCGAATTTGAGCAACAGG	900					

Qy	901	GTAAATATGCAAGTATGAATTCGTATAATAATAATGACCTTCGTTTCGCTATATAGAG	960
Db	901	GTAAATATGCAAGTATGAATTCGTATAATAATAATGACCTTCGTTTCGCTATATAGAG	960
Qy	961	ACTGCGGTATTCGAAAGCCGCACTACTACTTTCTAGAAACAACTTACAATTTTTAGC	1020
Db	961	ACTGCGGTATTCGAAAGCCGCACTACTACTTTCTAGAAACAACTTACAATTTTTAGC	1020
Qy	1021	ACTTCATCAGATGGAGTGTCTACTAGCATATAGCTTACTTGGCGGGGGGCAACAATTCAA	1080
Db	1021	ACTTCATCAGATGGAGTGTCTACTAGCATATAGCTTACTTGGCGGGGGGCAACAATTCAA	1080
Qy	1081	TCTCGGCCAATAGAGCGGATTAATACCTCAACGATGSGTCTACCAATACTTCTATT	1140
Db	1081	TCTCGGCCAATAGAGCGGATTAATACCTCAACGATGSGTCTACCAATACTTCTATT	1140
Qy	1141	AATCCTGTAAAGATTATCAITCTCTCGAGAGCTATATTGGAGTGAATCATATGACAGGA	1200
Db	1141	AATCCTGTAAAGATTATCAITCTCTCGAGAGCTATATTGGAGTGAATCATATGACAGGA	1200
Qy	1201	GTGCTTCTATGGGGAATTTACCTTGAACCTTATTCATGGTGTCCCTACTGTAGATTTAAT	1260
Db	1201	GTGCTTCTATGGGGAATTTACCTTGAACCTTATTCATGGTGTCCCTACTGTAGATTTAAT	1260
Qy	1261	TTTAGGAAACCTCAGAAATACTTTTGAAGAGGTACTGCTAACTATAGTCAACCTATGAG	1320
Db	1261	TTTAGGAAACCTCAGAAATACTTTTGAAGAGGTACTGCTAACTATAGTCAACCTATGAG	1320
Qy	1321	TCACCTGGGCTTCAATTTAAAGAAATCAGAAATCGAAATACCAACAGAAACAAACAGAACGA	1380
Db	1321	TCACCTGGGCTTCAATTTAAAGAAATCAGAAATCGAAATACCAACAGAAACAAACAGAACGA	1380
Qy	1381	CGAAATTTGAATCATATAGTCAATAGTTATCTCACATAGGGCTCATTTTCAATCTAGG	1440
Db	1381	CGAAATTTGAATCATATAGTCAATAGTTATCTCACATAGGGCTCATTTTCAATCTAGG	1440
Qy	1441	GTGCATGTACCAAGTATATCTTGGACGCAACCGTAGTGCAGATCGTACAAATACCATTAGT	1500
Db	1441	GTGCATGTACCAAGTATATCTTGGACGCAACCGTAGTGCAGATCGTACAAATACCATTAGT	1500
Qy	1501	TCAGATAGCAATACACAAATACCAATTTGGTAAATTCATTAACCTTAATTCAGGTACCTCT	1560
Db	1501	TCAGATAGCAATACACAAATACCAATTTGGTAAATTCATTAACCTTAATTCAGGTACCTCT	1560
Qy	1561	GTAGTCAGTGGCCAGGATTTTACAGAGGGGATTAATCCGAACTAACCGTTAATGGTAGT	1620
Db	1561	GTAGTCAGTGGCCAGGATTTTACAGAGGGGATTAATCCGAACTAACCGTTAATGGTAGT	1620
Qy	1621	GTACTAAGTATGGTCTTAAATTTTAAATTAATACATCAATACAGCGGTATCGCGTGAGAGTT	1680
Db	1621	GTACTAAGTATGGTCTTAAATTTTAAATTAATACATCAATACAGCGGTATCGCGTGAGAGTT	1680
Qy	1681	CGTTATGCTGTTCTCAAAACAATGGTCTGAGGGTAACTGTGCGAGGGAGTACTACTTTT	1740
Db	1681	CGTTATGCTGTTCTCAAAACAATGGTCTGAGGGTAACTGTGCGAGGGAGTACTACTTTT	1740
Qy	1741	GATCAAGGATTCCTAGTACTATGAGTGCATAATGAGTCTTTGACATCTCAATCATTTAGA	1800
Db	1741	GATCAAGGATTCCTAGTACTATGAGTGCATAATGAGTCTTTGACATCTCAATCATTTAGA	1800
Qy	1801	TTTTCAGAAATTTCTCTAGTATTTAGTGCATCTGGCAGTCAAACTGCTGGAATAAGTATA	1860
Db	1801	TTTTCAGAAATTTCTCTAGTATTTAGTGCATCTGGCAGTCAAACTGCTGGAATAAGTATA	1860
Qy	1861	AGTAATAATGCGAGGTAGACAAAGCTTTTCACTTTTGAATAAATTCGAATTCATTCCTACT	1920
Db	1861	AGTAATAATGCGAGGTAGACAAAGCTTTTCACTTTTGAATAAATTCGAATTCATTCCTACT	1920
Qy	1921	GCAACCTTCGAGCAGAACTACGATTTTAGAAAGGGCGCAAGAGCGGTGATGCTCTGTTT	1980
Db	1921	GCAACCTTCGAGCAGAACTACGATTTTAGAAAGGGCGCAAGAGCGGTGATGCTCTGTTT	1980

```
QY 1981 ACTAATACGAATCCAAAGATTGAAAAAGATGTGACAGATTATCATATTTGATCAAGTA 2040
Db 1981 ACTAATACGAATCCAAAGATTGAAAAAGATGTGACAGATTATCATATTTGATCAAGTA 2040
QY 2041 TCCAAATTTAGTGGCGTCTTTATCGAGTGAATTTCTGCTTAGATGAAAAGAGAGAAATTAATT 2100
Db 2041 TCCAAATTTAGTGGCGTCTTTATCGAGTGAATTTCTGCTTAGATGAAAAGAGAGAAATTAATT 2100
QY 2101 GAGAAAGTGAATATGCGAAACGACTCAGTGTAGTGAAGAAACCTTACTCCAGATCCAAAC 2160
Db 2101 GAGAAAGTGAATATGCGAAACGACTCAGTGTAGTGAAGAAACCTTACTCCAGATCCAAAC 2160
QY 2161 TTCACATCCATCAATAGCAACAGACTTCATATCTACTAATGAGCAATCGAATTTCCACA 2220
Db 2161 TTCACATCCATCAATAGCAACAGACTTCATATCTACTAATGAGCAATCGAATTTCCACA 2220
QY 2221 TCTATCCATGAACCAATCTGAACATGGATGGTGGGAAAGTGAGAACATTACAATCCAGGAA 2280
Db 2221 TCTATCCATGAACCAATCTGAACATGGATGGTGGGAAAGTGAGAACATTACAATCCAGGAA 2280
QY 2281 GGAATGACGTATTTAAAGAGAAATTAAGTACACTACCGGGGACTTTTAATGAGTGTAT 2340
Db 2281 GGAATGACGTATTTAAAGAGAAATTAAGTACACTACCGGGGACTTTTAATGAGTGTAT 2340
QY 2341 CCGACGTATTTATCAAAATAGGAGAGTCCGAATTAAGCTTATATCTCGTACCAA 2400
Db 2341 CCGACGTATTTATCAAAATAGGAGAGTCCGAATTAAGCTTATATCTCGTACCAA 2400
QY 2401 TTAAGAGGTATATTGAAGATAGTCAAGATTTAGAGATATATTGATTCGTATAATGCG 2460
Db 2401 TTAAGAGGTATATTGAAGATAGTCAAGATTTAGAGATATATTGATTCGTATAATGCG 2460
QY 2461 AAACATGAACATTGGATGTTCCAGGTAACCGATCCGATGCGCGCTTTCAAGTGAAGC 2520
Db 2461 AAACATGAACATTGGATGTTCCAGGTAACCGATCCGATGCGCGCTTTCAAGTGAAGC 2520
QY 2521 CCAATCGGAGGTCCGAGAACCGAATCGATCGCACCAATTGGAATGGAATCCGTAT 2580
Db 2521 CCAATCGGAGGTCCGAGAACCGAATCGATCGCACCAATTGGAATGGAATCCGTAT 2580
QY 2581 CTAGATTGTTCTTGACAGATGAGAAAAATGTGCGCATCATCCCATCATTTCTCTTTG 2640
Db 2581 CTAGATTGTTCTTGACAGATGAGAAAAATGTGCGCATCATCCCATCATTTCTCTTTG 2640
QY 2641 GATATTGATATTGATGCACAGACTTGCATGAGAAATCTAGGCGTGTGGGTGATTTCAAG 2700
Db 2641 GATATTGATATTGATGCACAGACTTGCATGAGAAATCTAGGCGTGTGGGTGATTTCAAG 2700
QY 2701 ATTAAGACGAGAGAGTCAATGCAAGACTAGGGAATCTGGAATTTATGGAAGAGAAACCA 2760
Db 2701 ATTAAGACGAGAGAGTCAATGCAAGACTAGGGAATCTGGAATTTATGGAAGAGAAACCA 2760
QY 2761 TTATTAGGAGACGACTGCTCGTGTGAAGAGGACGAGAAAAATGGAGACAAACGT 2820
Db 2761 TTATTAGGAGACGACTGCTCGTGTGAAGAGGACGAGAAAAATGGAGACAAACGT 2820
QY 2821 GAAAAACTACAATTTGAAAAACCAACAGTATATACAGAGGCAAAAGAGCTGTGGATGCT 2880
Db 2821 GAAAAACTACAATTTGAAAAACCAACAGTATATACAGAGGCAAAAGAGCTGTGGATGCT 2880
QY 2881 TTATTGTAGATTCTCAATTAATATAGATTACAGCGGATCAAAACATTCGCGATGATTCAT 2940
Db 2881 TTATTGTAGATTCTCAATTAATATAGATTACAGCGGATCAAAACATTCGCGATGATTCAT 2940
QY 2941 GCGGAGATAAATCTGTTTCAATTCGAATTCGAGAGCTTATCTGCAGAAATATCTGTATC 3000
Db 2941 GCGGAGATAAATCTGTTTCAATTCGAATTCGAGAGCTTATCTGCAGAAATATCTGTATC 3000
QY 3001 CCGGCTGTAATGCGGAAATTTTGAAGAAATTAAGAGTTCGCAATTCACCTGCAATCTCC 3060
Db 3001 CCGGCTGTAATGCGGAAATTTTGAAGAAATTAAGAGTTCGCAATTCACCTGCAATCTCC 3060
QY 3061 CTATACGATCGAGAAATGTCTGTTAAAAATGGTGAATTTTAAATGGAATGATGATGCTG 3120
```

```
Db 3061 CTATACGATCGAGAAATGTCTGTTAAAAATGGTGAATTTTAAATGGAATGATGATGCTG 3120
QY 3121 AATGTTAAAAAGGCGATGTAGATGTACAACAGAGCCATCACCGTTCGTCTGTTGTTATCCCA 3180
Db 3121 AATGTTAAAAAGGCGATGTAGATGTACAACAGAGCCATCACCGTTCGTCTGTTGTTATCCCA 3180
QY 3181 GAATGGGAGCAGAAAGTGTCACAAGCAGTTCCGCTCTGTCCGGGCGTGGCTATATCTC 3240
Db 3181 GAATGGGAGCAGAAAGTGTCACAAGCAGTTCCGCTCTGTCCGGGCGTGGCTATATCTC 3240
QY 3241 CGTGTCAACGCGTACAAAGAGGATATGAGAGGGTGTGTAAACGATCCCATGAATCGAG 3300
Db 3241 CGTGTCAACGCGTACAAAGAGGATATGAGAGGGTGTGTAAACGATCCCATGAATCGAG 3300
QY 3301 AACAAATACAGACGAACCTAAATTTAAAACTGTGAAGAGGAGAGTGTATCCAAACGGAT 3360
Db 3301 AACAAATACAGACGAACCTAAATTTAAAACTGTGAAGAGGAGAGTGTATCCAAACGGAT 3360
QY 3361 ACAGGAACGTGTAAATGATATATCTGCACCAAGGATACAGAGTATGTAATTCCTGTAAT 3420
Db 3361 ACAGGAACGTGTAAATGATATATCTGCACCAAGGATACAGAGTATGTAATTCCTGTAAT 3420
QY 3421 GCTGGATATGAGGATGATATGAGTGTGATCTACAGCATCTGTTAATTAACAACCGACT 3480
Db 3421 GCTGGATATGAGGATGATATGAGTGTGATCTACAGCATCTGTTAATTAACAACCGACT 3480
QY 3481 TATGAAGAGAAACGCTATACAGATGTACGAAGAGATAATCATTTGTAATATGACAGAGG 3540
Db 3481 TATGAAGAGAAACGCTATACAGATGTACGAAGAGATAATCATTTGTAATATGACAGAGG 3540
QY 3541 TATGTGAATATCCACCACTACAGCTGGTTATATGACAAAAGAAATAGAAATTTCCCA 3600
Db 3541 TATGTGAATATCCACCACTACAGCTGGTTATATGACAAAAGAAATAGAAATTTCCCA 3600
QY 3601 GAAACCGATAAGCTATGGAATTTGAGATTGGAAGAACCGAAGGAGTTTATTGTAGACAGC 3660
Db 3601 GAAACCGATAAGCTATGGAATTTGAGATTGGAAGAACCGAAGGAGTTTATTGTAGACAGC 3660
QY 3661 GTGGAATTAATCTCTTATGAGGAAATAG 3687
Db 3661 GTGGAATTAATCTCTTATGAGGAAATAG 3687
```

RESULT 2

AR359382 LOCUS AR359382 4173 bp DNA linear PAT 17-AUG-2003

DEFINITION Sequence 37 from patent US 6593293.

ACCESSION AR359382

VERSION AR359382.1 GI:33765692

KEYWORDS

SOURCE Unknown.

ORGANISM

Unclassified.

REFERENCE 1 (bases 1 to 4173)

AUTHORS Baum,J.A., Chu,C.-R., Donovan,W.P., Gilmer,A.J. and Rupar,M.J.

TITLE Lepidopteran-active Bacillus thuringiensis .delta.-endotoxin

compositions and methods of use

JOURNAL Patent: US 6593293-A 37 15-JUL-2003;

FEATURES Location/Qualifiers

source 1. 4173

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 99.4%; Score 3666.2; DB 6; Length 4173;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 3674; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 TTGACTTCAATAGGAAAAATGAGAAATTAATAATGCTTTTATCGATTTCAGGCTGTA 60
Db 1 TTGACTTCAATAGGAAAAATGAGAAATTAATAATGCTTTTATCGATTTCAGGCTGTA 60

Qy 61 TCGAATCATTTCCACACAAATGGATCTATACACAGATGCTCGTATGTAGGATTTCTTTGTG 120
 Db |||||
 Qy 61 TCGAATCATTTCCACACAAATGGATCTATACACAGATGCTCGTATGTAGGATTTCTTTGTG 120
 Db |||||
 Qy 121 ATAGCCAGGGGATATATCAATCCACTTGTGTAGCGCATCAACAGTCCAAACGGGTATT 180
 Db |||||
 Qy 121 ATAGCCAGGGGATATATCAATCCACTTGTGTAGCGCATCAACAGTCCAAACGGGTATT 180
 Db |||||
 Qy 181 AACATAGCTGTAGAACTACTAGGTGTATAGGCGTACCGTTTGTGTCAGCAAAATAGCTAGT 240
 Db |||||
 Qy 181 AACATAGCTGTAGAACTACTAGGTGTATAGGCGTACCGTTTGTGTCAGCAAAATAGCTAGT 240
 Db |||||
 Qy 241 TTTTATAGTTTTCTTTGTGTGAATTAATGCGCCGCGCAGAGATCAGTGGGAAATTTTC 300
 Db |||||
 Qy 241 TTTTATAGTTTTCTTTGTGTGAATTAATGCGCCGCGCAGAGATCAGTGGGAAATTTTC 300
 Db |||||
 Qy 301 CTAGAACATGTGCAACAACTTATAAATCAACAAATTAACAGAAATGCTAGGAATACGGCA 360
 Db |||||
 Qy 301 CTAGAACATGTGCAACAACTTATAAATCAACAAATTAACAGAAATGCTAGGAATACGGCA 360
 Db |||||
 Qy 361 CTTGCTCGATTACAGGTTTAGGAGATTCCTTTAGAGCCCTATCAACAGTCACTTGAAGAT 420
 Db |||||
 Qy 361 CTTGCTCGATTACAGGTTTAGGAGATTCCTTTAGAGCCCTATCAACAGTCACTTGAAGAT 420
 Db |||||
 Qy 421 TGGCTAGAAAACCGTGATGATGCAAGAAACGAGAGTGTCTTTATATACCCAAATATAGCC 480
 Db |||||
 Qy 421 TGGCTAGAAAACCGTGATGATGCAAGAAACGAGAGTGTCTTTATATACCCAAATATAGCC 480
 Db |||||
 Qy 481 TTAGAACTTGAATTTCTTAATGCGATGCGCGCTTTTCGCAATTAAGAACCAAGAGTTCCA 540
 Db |||||
 Qy 481 TTAGAACTTGAATTTCTTAATGCGATGCGCGCTTTTCGCAATTAAGAACCAAGAGTTCCA 540
 Db |||||
 Qy 541 TTATTAATGATATGCTCAAGCTGCAAAATTTACACCTATTATTATTAGAGATGCTCTCT 600
 Db |||||
 Qy 541 TTATTAATGATATGCTCAAGCTGCAAAATTTACACCTATTATTATTAGAGATGCTCTCT 600
 Db |||||
 Qy 601 CTTTTTGGTATGAAATTTGGGCTTACATCGCAGAAATTCAGCTTTATATAGCGCCCAA 660
 Db |||||
 Qy 601 CTTTTTGGTATGAAATTTGGGCTTACATCGCAGAAATTCAGCTTTATATAGCGCCCAA 660
 Db |||||
 Qy 661 GTGGAACAAACGAGAGATTTCCGACTATGCTAGAAATGGTATATATACAGTCTTAAT 720
 Db |||||
 Qy 661 GTGGAACAAACGAGAGATTTCCGACTATGCTAGAAATGGTATATATACAGTCTTAAT 720
 Db |||||
 Qy 721 AGCTTGAGAGGCAAAATGCGCAAGTTGGGTGCTTATTAATCAATTTCCGTAGAGATCTA 780
 Db |||||
 Qy 721 AGCTTGAGAGGCAAAATGCGCAAGTTGGGTGCTTATTAATCAATTTCCGTAGAGATCTA 780
 Db |||||
 Qy 781 ACGTTAGGGTATFAGATCTAGTGGCACTATTCCTCAAGCTATGACACTCGCACTTATCCA 840
 Db |||||
 Qy 781 ACGTTAGGGTATFAGATCTAGTGGCACTATTCCTCAAGCTATGACACTCGCACTTATCCA 840
 Db |||||
 Qy 841 ATAAATACGAGTCTCAGTTTAAACAGGGAAGTTTATACAGCCCAATGAGGCAACAGGG 900
 Db |||||
 Qy 841 ATAAATACGAGTCTCAGTTTAAACAGGGAAGTTTATACAGCCCAATGAGGCAACAGGG 900
 Db |||||
 Qy 901 GTAATATGCAAGTATGAAATGGTATTAATAAATGCACCTTCGTTTTCGGCTATAGAG 960
 Db |||||
 Qy 901 GTAATATGCAAGTATGAAATGGTATTAATAAATGCACCTTCGTTTTCGGCTATAGAG 960
 Db |||||
 Qy 961 ACTGCGGTTATCCGAAACCGCGATCTACTTGAATTTCTAGAAACAACTTAATTTTTCAGC 1020
 Db |||||
 Qy 961 ACTGCGGTTATCCGAAACCGCGATCTACTTGAATTTCTAGAAACAACTTAATTTTTCAGC 1020
 Db |||||
 Qy 1021 ACTTCATCAGATGAGTGTCTAGGCAATGCTTACTTGGCGGGGCGCACAAATTTCAA 1080
 Db |||||
 Qy 1021 ACTTCATCAGATGAGTGTCTAGGCAATGCTTACTTGGCGGGGCGCACAAATTTCAA 1080
 Db |||||
 Qy 1081 TCTCGGCCAATAGAGGCGGATTAATACCTCAACGATGGGTCTAGCAATATCTTCTATT 1140
 Db |||||
 Qy 1081 TCTCGGCCAATAGAGGCGGATTAATACCTCAACGATGGGTCTAGCAATATCTTCTATT 1140
 Db |||||
 Qy 1141 AATCCTGTAAGATTATCATTTCTCTCTCGAGACGTATATTGGAGTGAATCATATGCAAGGA 1200
 Db |||||

Db 1141 AATCCTGTAAGATTATCATTTCTCTCTCGAGACGTATATTGGACTGAATCATATGCAGGA 1200
 Qy |||||
 Db 1201 GTGCTTCTATGGGGAATTTTACCTTTGAACTTATTCATGGTGTCCCTACTGTGTAGATTTAAT 1260
 Qy |||||
 Db 1201 GTGCTTCTATGGGGAATTTTACCTTTGAACTTATTCATGGTGTCCCTACTGTGTAGATTTAAT 1260
 Qy |||||
 Db 1261 TTTAGGAACCTCTCAGAAATACTTTTGAAGAGGTACTGCTAACTATATAGTCAACCTATGAG 1320
 Qy |||||
 Db 1261 TTTAGGAACCTCTCAGAAATACTTTTGAAGAGGTACTGCTAACTATATAGTCAACCTATGAG 1320
 Qy |||||
 Db 1321 TCACCTGGGCTTCAATTTAAAGATTCAGAAATCTGAATTTACCAACAGAAACAAACAGAACGA 1380
 Qy |||||
 Db 1321 TCACCTGGGCTTCAATTTAAAGATTCAGAAATCTGAATTTACCAACAGAAACAAACAGAACGA 1380
 Qy |||||
 Db 1381 CCAAAATATGAATCATATAGTCAATAGTTTATCTCAGATAGGGCTCAITTCACAACTTAGG 1440
 Qy |||||
 Db 1381 CCAAAATATGAATCATATAGTCAATAGTTTATCTCAGATAGGGCTCAITTCACAACTTAGG 1440
 Qy |||||
 Db 1441 GTGCATGTACAGTATATTTCTTGGAGCGCACCGTAGTGCAGATCGTACAAATACCATTAGT 1500
 Qy |||||
 Db 1441 GTGCATGTACAGTATATTTCTTGGAGCGCACCGTAGTGCAGATCGTACAAATACCATTAGT 1500
 Qy |||||
 Db 1501 TCAGATAGCAATACAAATACCAATTTGTTGTTAAATCATTTCAACCTTAATTCAGGTACTCT 1560
 Qy |||||
 Db 1501 TCAGATAGCAATACAAATACCAATTTGTTGTTAAATCATTTCAACCTTAATTCAGGTACTCT 1560
 Qy |||||
 Db 1561 GTAGTCAGTGGCCAGGATTTACAGAGGGGATATTAATCCGAACTAACGTTTAATGGTAGT 1620
 Qy |||||
 Db 1561 GTAGTCAGTGGCCAGGATTTACAGAGGGGATATTAATCCGAACTAACGTTTAATGGTAGT 1620
 Qy |||||
 Db 1621 GTACTAAGTATGGTCTTAATTTTAAATPACATCATTTACAGCGGTATCCGCTGAGAGTT 1680
 Qy |||||
 Db 1621 GTACTAAGTATGGTCTTAATTTTAAATPACATCATTTACAGCGGTATCCGCTGAGAGTT 1680
 Qy |||||
 Db 1681 CGTTATGCTGCTTCTCAACAAATGGTCTTGAGGGTAACTGTCCGAGGAGTACTACTTTTT 1740
 Qy |||||
 Db 1681 CGTTATGCTGCTTCTCAACAAATGGTCTTGAGGGTAACTGTCCGAGGAGTACTACTTTTT 1740
 Qy |||||
 Db 1741 GATCAAGGATTCCTAGTACTATGAGTGCATAATGAGTCTTTGACATCTCAATCATTTTGA 1800
 Qy |||||
 Db 1741 GATCAAGGATTCCTAGTACTATGAGTGCATAATGAGTCTTTGACATCTCAATCATTTGA 1800
 Qy |||||
 Db 1801 TTTGCAAGATTTCTGTAGGTATTTAGTGCATTCGGCAGTCAAACTGCTGGAATAAGTATA 1860
 Qy |||||
 Db 1801 TTTGCAAGATTTCTGTAGGTATTTAGTGCATTCGGCAGTCAAACTGCTGGAATAAGTATA 1860
 Qy |||||
 Db 1861 AGTAATAATGCAAGTATAGCAAACTGTTTCACTTTGATAAATTTGAATTCATTTCAATTTACT 1920
 Qy |||||
 Db 1861 AGTAATAATGCAAGTATAGCAAACTGTTTCACTTTGATAAATTTGAATTCATTTCAATTTACT 1920
 Qy |||||
 Db 1921 GCAACCTTTCGAAGCAGAAATACGATTTTAGAAAGGGCGCAAGCGGTGAATGCTCTGTTTT 1980
 Qy |||||
 Db 1921 GCAACCTTTCGAAGCAGAAATACGATTTTAGAAAGGGCGCAAGCGGTGAATGCTCTGTTTT 1980
 Qy |||||
 Db 1981 ACTAATAACGAATCCAAAGAGATTTGAAACAGATGTGACAGATTTATCATATTGATCAAGTA 2040
 Qy |||||
 Db 1981 ACTAATAACGAATCCAAAGAGATTTGAAACAGATGTGACAGATTTATCATATTGATCAAGTA 2040
 Qy |||||
 Db 2041 TCCAAATTTAGTGGGCTGTTTATCGGATGAATTTCTGCTTAGATGAAAAGAGAGAAATTTACT 2100
 Qy |||||
 Db 2041 TCCAAATTTAGTGGGCTGTTTATCGGATGAATTTCTGCTTAGATGAAAAGAGAGAAATTTACT 2100
 Qy |||||
 Db 2101 GAGAAAGTGAATAATGCGAAACGACTCAGTGTGTAAGAAACCTTACTCCAAAGATCCAAAC 2160
 Qy |||||
 Db 2101 GAGAAAGTGAATAATGCGAAACGACTCAGTGTGTAAGAAACCTTACTCCAAAGATCCAAAC 2160
 Qy |||||
 Db 2161 TTCAATCCATCAATTAAGCAACAGACTTTCATATCTAATGAGCAATCGAATTTTCA 2220
 Qy |||||
 Db 2161 TTCAATCCATCAATTAAGCAACAGACTTTCATATCTAATGAGCAATCGAATTTTCA 2220
 Qy |||||
 Db 2221 TCTATCCATGAACAACTGNAACATGATGTTGGGAGTGAGAACATTAACAATCCAGGAA 2280
 Qy |||||

```

Db      2221  TCTATCCATGAACAATCTGAACATGGATGGTGGGAAAGTGAGAACATTTACAATCCAGAA 2280
Qy      2281  GGAATGACGTATTTAAAGAAATTACGTCACTACCGGGACCTTTTAATGAGTGTAT 2340
Db      2281  GGAATGACGTATTTAAAGAAATTACGTCACTACCGGGACCTTTTAATGAGTGTAT 2340
Qy      2341  CCGACGTATTTATATCAAAAATAGGAGAGTCGGAATTTAAAGCTTATATCTCGTACCAA 2400
Db      2341  CCGACGTATTTATATCAAAAATAGGAGAGTCGGAATTTAAAGCTTATATCTCGTACCAA 2400
Qy      2401  TTAAGAGGTATATTTGAAGTAGTCAAGATTTAGAGATATATTTGATTCGTTATAATGCG 2460
Db      2401  TTAAGAGGTATATTTGAAGTAGTCAAGATTTAGAGATATATTTGATTCGTTATAATGCG 2460
Qy      2461  AAACATGAACATTGGATGTTCCAGGTACCGAGTCGATGCGCGCTTTTCAGTTGAAAGC 2520
Db      2461  AAACATGAACATTGGATGTTCCAGGTACCGAGTCGATGCGCGCTTTTCAGTTGAAAGC 2520
Qy      2521  CCAATCGGAAGTCCGAGAACCGAATCGATGCGCACCAATTTTGAATGGAATCCTGAT 2580
Db      2521  CCAATCGGAAGTCCGAGAACCGAATCGATGCGCACCAATTTTGAATGGAATCCTGAT 2580
Qy      2581  CTAGATTGTTCTCGACAGATGGAAGAAATGTCCGCATCATTCGCCATCATTTCTCTTTG 2640
Db      2581  CTAGATTGTTCTCGACAGATGGAAGAAATGTCCGCATCATTCGCCATCATTTCTCTTTG 2640
Qy      2641  GATATTGATATTTGGATGCACAGACTTCGATGAGAAATCTAGCGGTGGTGGTATTCAG 2700
Db      2641  GATATTGATATTTGGATGCACAGACTTCGATGAGAAATCTAGCGGTGGTGGTATTCAG 2700
Qy      2701  ATTAAGACGAGAGAGTGCATGAAAGTACGAGGATCTGGAAATTTATTTGAAGAGAACCA 2760
Db      2701  ATTAAGACGAGAGAGTGCATGAAAGTACGAGGATCTGGAAATTTATTTGAAGAGAACCA 2760
Qy      2761  TTATTAGGAGACGACGTCTCGTGTGACAGACGACAGAGAAATGAGAGACAAACGT 2820
Db      2761  TTATTAGGAGACGACGTCTCGTGTGAGAGAGACAGAGAAATGAGAGACAAACGT 2820
Qy      2821  GAAAACTACAATTTGAAAAACAAACGAGTATATACAGAGGCAAAAGCTGTGGATGCT 2880
Db      2821  GAAAACTACAATTTGAAAAACAAACGAGTATATACAGAGGCAAAAGCTGTGGATGCT 2880
Qy      2881  TTATTGTAGATCTCAATATATAGATTACAGCGGATCAACACATTTGGATGATTCAT 2940
Db      2881  TTATTGTAGATCTCAATATATAGATTACAGCGGATCAACACATTTGGATGATTCAT 2940
Qy      2941  GCGGCAGATAAATGTTTCAATTCGAATTCGAGAGCTTATCTGTGAGAAATATCTGTTATC 3000
Db      2941  GCGGCAGATAAATGTTTCAATTCGAATTCGAGAGCTTATCTGTGAGAAATATCTGTTATC 3000
Qy      3001  CCGGGTGAATGCGGAAATTTTGAAGAAATTTGAAGGTCGATTAATCACTGCAATCTCC 3060
Db      3001  CCGGGTGAATGCGGAAATTTTGAAGAAATTTGAAGGTCGATTAATCACTGCAATCTCC 3060
Qy      3061  CTATACGATGCGGAATGCTGTTAAATGTTGATTTTAAATGATTTAGATGATGCTGG 3120
Db      3061  CTATACGATGCGGAATGCTGTTAAATGTTGATTTTAAATGATTTAGATGATGCTGG 3120
Qy      3121  AATGTAAGGGCATGTAGATGTACACAGAGCCATCACCGTCTCTGTTGTTATCCCA 3180
Db      3121  AATGTAAGGGCATGTAGATGTACACAGAGCCATCACCGTCTCTGTTGTTATCCCA 3180
Qy      3181  GAATGGGAAGCAGAAAGTGTCAACAGCAGTTCGCTGTCTGCGGGCGTGGCTATATCCTC 3240
Db      3181  GAATGGGAAGCAGAAAGTGTCAACAGCAGTTCGCTGTCTGCGGGCGTGGCTATATCCTC 3240
Qy      3241  CGTGTACAGCGTACAAAGAGGATATGGAGAGGTTGTGTAAACGATCCATGAATCCAG 3300
Db      3241  CGTGTACAGCGTACAAAGAGGATATGGAGAGGTTGTGTAAACGATCCATGAATCCAG 3300
Qy      3301  AACAAATCAGACGAATCAAAATTTAAATACTGTGAAGAGGAGTGTATCCAAAGCAT 3360
Db      3301  AACAAATCAGACGAATCAAAATTTAAATACTGTGAAGAGGAGTGTATCCAAAGCAT 3360

```

```

Qy      3361  ACAGGACGTTGTAATGATTTATCTGACACCAAGGTACAGCATCTATCTAATTCCTGTAAT 3420
Db      3361  ACAGGACGTTGTAATGATTTATCTGACACCAAGGTACAGCATCTAATTCCTGTAAT 3420
Qy      3421  GCTGGATATCAGGATGCATATGAAGTTGATCTACAGCATCTGTTAATTTACAAACCGACT 3480
Db      3421  GCTGGATATCAGGATGCATATGAAGTTGATCTACAGCATCTGTTAATTTACAAACCGACT 3480
Qy      3481  TATGAAGAAACGATATACAGATGATGACGAAGATTAATCATTTGTGTAATGACAGAGGG 3540
Db      3481  TATGAAGAAACGATATACAGATGATGACGAAGATTAATCATTTGTGTAATGACAGAGGG 3540
Qy      3541  TATGTGAATTTATCCACACTACCGCTGTTATATGACAAAGAAATAGATATCTTCCCA 3600
Db      3541  TATGTGAATTTATCCACACTACCGCTGTTATGTCGACAAAGAAATAGATATTTTCCCA 3600
Qy      3601  GAAACCGATTAAGTATGATTTGAGATTGGAGAAACGGAAGGAAGTTTATTTGTAGACAGC 3660
Db      3601  GAAACCGATTAAGTATGATTTGAGATTGGAGAAACGGAAGGAAGTTTATTTGTAGACAGC 3660
Qy      3661  GTGGAATTAATCTCTCTTATGGAGGAATAG 3687
Db      3661  ATAGAATTAATCTCTCTTATGGAGGAATAG 3687

```

RESULT 3

```

AX098667
LOCUS      AX098667               4173 bp      DNA      linear      PAT 02-APR-2001
DEFINITION Sequence 37 from Patent WO0119859.
ACCESSION AX098667
VERSION    AX098667.1 GI:13537912
KEYWORDS   .
SOURCE     Bacillus thuringiensis
            Bacillus thuringiensis
            Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
            cereus group.

```

REFERENCE

Baum, J.A., Chu, C.R., Donovan, W.P., Gilmer, A.J. and Rupar, M.J.
 Lepidopteran-active bacillus thuringiensis delta-endotoxin
 compositions and methods of use
 Patent: WO 0119859-A 37 22-MAR-2001;
 MONSANTO COMPANY (US)

FEATURES

```

source     Location/Qualifiers
            1..4173
            /organism="Bacillus thuringiensis"
            /mol_type="unassigned DNA"
            /db_xref="taxon:1428"
            1..3687
            /note="unnamed protein product"
            /codon_start=1
            /transl_table=11
            /protein_id="CAC35766.1"
            /db_xref="GI:13537913"

```

CDS

/translation="MISNRKNEIINALSIPAVSHSTQMDLSPDARIEDSLCIABG
 NNINPLVSTAVTQGINIAGILVGLVGPAGQIASFYSLVGLMWRGRDQWEIFLE
 HVEQLNQITENARNTALRQLGDSFRAYQQSLDLENDLDRVTSVLYTQYTA
 LEUFLNMLPLFAIRNOEVLPLAVYQAANLHLLLDLADSLFGSEGLTQOEIORYYE
 ROVEQTRDYSDYCVENYNTGLNSLRGTNAASWYRNVQFRDLTLGLVLDLVALPSPYDT
 RYTPNTSAQLTREVYTDALGATGYNWASNNWNNRNPSPSAIETAVIRPSLDLLE
 QLTFTSSRSWTSATRYWRGHTIQSRPIGGGLNTSTHGSTNTSINPRLSPFRSDV
 YMTSEYGLLWGLIYLEPIHGVTYRFRNFNPQNTFERTGNTSQPYESPLQKDE
 TELPPTETTPPNYSYSHRLSLGLISQSRVHPVTVSWTHRSADRTWTISDSITQIP
 LVKSFNLSGTSVSGPGTGGDIIRTNVNGSVLSMGLNFNTSLQRYRVRVAAQ
 TWLVRYVSGSTTFDQGFSTMSANBSLTSQSFRFAEPVGISASGQTAGISLNNNA
 GRQTFHDKIEFIPITATPEAYDLERAQAVNALFTNTPRLKTDVTDHYHDQVSN
 LVAQLDEFCLEKDEKLEKVKYAKSLDERNLQDPNFTSINKPQDFISLNEQSNFT
 SIHQSEHGWGWNENITIQEONDVPKENYVTLPGTNECYPTLYLQKIGSELKAYTR
 YQLRQYTEDSQDLEIYLIRYNKXHELDVPGTESVMPLSVESPIGRGEPNRCAPHE
 WNPDLDCSRDEKCAHSHHFLDIDIGCTDLHENLGVWVFKITQOEHARGNLE
 FISEKPLLGALSRVRAEKWRDKREKLETKRVYVTEAKEAVDALFTDQSYNRLQA
 DTNIGMHAADKLIVRIEATLSVLPGNAAIFEEELGRIITLISLDLDRNVVKN
 GDFNGLACNVNKHVDVQQSHRSVLVPIPEWEAEVSQAVRVCPRGVIILRVTAIKSG

YGGCVTHIEIENNTDELKFKNCEBEEVYPTDGTGNDYTHAGHTAACNSRNAGYBDA
YEVDTTASVNYKYTEBETTYDVRDHNHCEYDRGYNNYPLPAGYVTKLEYFPETDK
WVIEIGTEGTFIVDSIELLMEE"

ORIGIN

Query Match		99.4%; Score 3666.2; DB 6; Length 4173;
Best Local Similarity		99.6%; Pred. No. 0;
Matches 3674; Conservative 0; Mismatches 13; Indels 0; Gaps 0;		
Qy	1	TTGACTTCAAAATAGGAAAATGAGAATGAATATATAATGCTTTATCGATTCCAGCTGTA 60
Db	1	TTGACTTCAAAATAGGAAAATGAGAATGAATATATAATGCTTTATCGATTCCAGCTGTA 60
Qy	61	TGCAATCATCCACACAAATGGATCTATCACAGATGCTGTAATGAGGATTTCTTTGTG 120
Db	61	TGCAATCATTTCCACACAAATGGATCTATCACAGATGCTGTAATGAGGATTTCTTTGTG 120
Qy	121	ATAGCCAGGGGAATATATCAATCCACTTTGTTAGCGCATCAACAGTCCAAACGGGTATT 180
Db	121	ATAGCCAGGGGAATATATCAATCCACTTTGTTAGCGCATCAACAGTCCAAACGGGTATT 180
Qy	181	AACATAGCTGTAGAACTACTAGGTGATATTAGGCGTACCGTTTCGTGACAAATAGCTAGT 240
Db	181	AACATAGCTGTAGAACTACTAGGTGATATTAGGCGTACCGTTTCGTGACAAATAGCTAGT 240
Qy	241	TTTTATAGTTTCTTGTGTTGTAATTATGCGCCGCGCGGAGAGATCAGTCGGGAAATTTTC 300
Db	241	TTTTATAGTTTCTTGTGTTGTAATTATGCGCCGCGCGGAGAGATCAGTCGGGAAATTTTC 300
Qy	301	CTAGAACATGTGCAACAACTTATAATCAACAAATAACAGAAATGCTAGGAATACGGCA 360
Db	301	CTAGAACATGTGCAACAACTTATAATCAACAAATAACAGAAATGCTAGGAATACGGCA 360
Qy	361	CTTGCTCGATTACAAGTTTAGAGATTCTTTAGAGCCCTATCAACAGTCACCTTGAAGAT 420
Db	361	CTTGCTCGATTACAAGTTTAGAGATTCTTTAGAGCCCTATCAACAGTCACCTTGAAGAT 420
Qy	421	TGGCTAGAAAACCGTGATGATGCAAGAACGAGAAGTGTCTTTATACCCCAATATAGCC 480
Db	421	TGGCTAGAAAACCGTGATGATGCAAGAACGAGAAGTGTCTTTATACCCCAATATAGCC 480
Qy	481	TTAGAACTTGATTTCTTAATGCGATGCGCTTTTGGCAATTAAGAACCAAGATTTCCA 540
Db	481	TTAGAACTTGATTTCTTAATGCGATGCGCTTTTGGCAATTAAGAACCAAGATTTCCA 540
Qy	541	TTATTATGTTATGCTCAAGTCGCAAAATTTACACCTATTATTATTGAGAGATGCTCT 600
Db	541	TTATTATGTTATGCTCAAGTCGCAAAATTTACACCTATTATTATTGAGAGATGCTCT 600
Qy	601	CTTTTGGTATGAAATTTGGGCTTACATCGCAGGAAATTCACGTTATTATGAGCGCCAA 660
Db	601	CTTTTGGTATGAAATTTGGGCTTACATCGCAGGAAATTCACGTTATTATGAGCGCCAA 660
Qy	661	GTGGAAACACGAGAGATTATTCGACTATGCGTAGAATGGTATATACAGTCTAAAT 720
Db	661	GTGGAAACACGAGAGATTATTCGACTATGCGTAGAATGGTATATACAGTCTAAAT 720
Qy	721	AGCTTGAGAGGACAAATGCGCAAGTTGGTGTGTTATATCAATTTCCGTAGAGATCTA 780
Db	721	AGCTTGAGAGGACAAATGCGCAAGTTGGTGTGTTATATCAATTTCCGTAGAGATCTA 780
Qy	781	ACGTTAGGGGATTTAGATCTAGTGGCACTATTCCTCAAGCTATGACATCTGCACTTATCCA 840
Db	781	ACGTTAGGGGATTTAGATCTAGTGGCACTATTCCTCAAGCTATGACATCTGCACTTATCCA 840
Qy	841	ATAAATACGAGTCTCAGTTTAAACAGGGAAGTTTATACAGCGCAATTTGGAGCAACGGG 900
Db	841	ATAAATACGAGTCTCAGTTTAAACAGGGAAGTTTATACAGCGCAATTTGGAGCAACGGG 900
Qy	901	GTAATATGCAAGTATGAATTTGGTATAATAATGACCTTCGTTTTCGGCTATAGAG 960
Db	901	GTAATATGCAAGTATGAATTTGGTATAATAATGACCTTCGTTTTCGGCTATAGAG 960

Qy	961	ACTGCGGTTATCCGAGCCGCACTACTACTAGGCATATGACTTCTGGCGGGGACACAATTCAA 1080
Db	961	ACTGCGGTTATCCGAGCCGCACTACTACTAGGCATATGACTTCTGGCGGGGACACAATTCAA 1020
Qy	1021	ACTTCATCAGATGGAGTGTCTACTAGGCATATGACTTCTGGCGGGGACACAATTCAA 1080
Db	1021	ACTTCATCAGATGGAGTGTCTACTAGGCATATGACTTCTGGCGGGGACACAATTCAA 1080
Qy	1081	TCTCGGCCAATAGCAGCGGATTAATATCCTCAACGCATGGGTCTACCAATATCTCTATT 1140
Db	1081	TCTCGGCCAATAGCAGCGGATTAATATCCTCAACGCATGGGTCTACCAATATCTCTATT 1140
Qy	1141	AATCTCTGTAAGATTATCACTTCTCTCGAGACGTATATTGGACTGAATCATATGAGGA 1200
Db	1141	AATCTCTGTAAGATTATCACTTCTCTCGAGACGTATATTGGACTGAATCATATGAGGA 1200
Qy	1201	GTGCTTCTATGGGGAATTTACCTTGAACCTTATTCATGGGTGTCCTACTGTTAGATTAA 1260
Db	1201	GTGCTTCTATGGGGAATTTACCTTGAACCTTATTCATGGGTGTCCTACTGTTAGATTAA 1260
Qy	1261	TTTAGGAACCTCAGAAATACCTTTTGAAGAGGTACTGCTAACTATAGTCAACCTATGAG 1320
Db	1261	TTTAGGAACCTCAGAAATACCTTTTGAAGAGGTACTGCTAACTATAGTCAACCTATGAG 1320
Qy	1321	TCACCTGGGCTTCAATTTAAAGATTGAGAACTGAATACCAACAGAAACAAACAGAACGA 1380
Db	1321	TCACCTGGGCTTCAATTTAAAGATTGAGAACTGAATACCAACAGAAACAAACAGAACGA 1380
Qy	1381	CGAAATATGCAATCATATAGTCACTAGTTATCTCATAGGCTCATTTTACCAATCTAGG 1440
Db	1381	CGAAATATGCAATCATATAGTCACTAGTTATCTCATAGGCTCATTTTACCAATCTAGG 1440
Qy	1441	GTGCATGTACCAAGTATATTTCTTGGAGCGCACCGTAGTGCAGATCGTACAAATACCAT 1500
Db	1441	GTGCATGTACCAAGTATATTTCTTGGAGCGCACCGTAGTGCAGATCGTACAAATACCAT 1500
Qy	1501	TCAGATAGATTAACAACAATACCATTTGGTAAATCAATCAATCCTTAATTCAGGTACTCT 1560
Db	1501	TCAGATAGATTAACAACAATACCATTTGGTAAATCAATCAATCCTTAATTCAGGTACTCT 1560
Qy	1561	GTAGTCAGTGGGCCAGGATTTACAGGAGGGATATATCCGAACTAACGTTAATGTTAGT 1620
Db	1561	GTAGTCAGTGGGCCAGGATTTACAGGAGGGATATATCCGAACTAACGTTAATGTTAGT 1620
Qy	1621	GTACTAAGTATGGGTCTTAATTTTAAATAATACATCAATCAGCGGTATCGCGTGAGATT 1680
Db	1621	GTACTAAGTATGGGTCTTAATTTTAAATAATACATCAATCAGCGGTATCGCGTGAGATT 1680
Qy	1681	CGTTATGCTGCTTCTCAAAATGCTGAGGGTAACTGTCGGAGGGAGTACTACTTTT 1740
Db	1681	CGTTATGCTGCTTCTCAAAATGCTGAGGGTAACTGTCGGAGGGAGTACTACTTTT 1740
Qy	1741	GATCAAGGATTCCTAGTACTATGAGTCAAAATGAGTCTTTTGACATCTCAATCATTTAGA 1800
Db	1741	GATCAAGGATTCCTAGTACTATGAGTCAAAATGAGTCTTTTGACATCTCAATCATTTAGA 1800
Qy	1801	TTTGCAAGATTTTCTCTAGGTATTTAGTGCATCTCGGAGTCAAACTGCTGGAATAAGTATA 1860
Db	1801	TTTGCAAGATTTTCTCTAGGTATTTAGTGCATCTCGGAGTCAAACTGCTGGAATAAGTATA 1860
Qy	1861	AGTAATATGCGAGTATAGCAAAAGGTTTCACTTTTGAATAAATGGAATTCATTCOAATTACT 1920
Db	1861	AGTAATATGCGAGTATAGCAAAAGGTTTCACTTTTGAATAAATGGAATTCATTCOAATTACT 1920
Qy	1921	GCAACCTTCGAGCAGAAATACGATTTTGAAGAGGGCGCAAGAGCGGTGATGCTCTGTTT 1980
Db	1921	GCAACCTTCGAGCAGAAATACGATTTTGAAGAGGGCGCAAGAGCGGTGATGCTCTGTTT 1980
Qy	1981	ACTAATACGAATCCAAAGAGATTGAAACAGATGTGACAGATTATCATATTGATCAAGTA 2040
Db	1981	ACTAATACGAATCCAAAGAGATTGAAACAGATGTGACAGATTATCATATTGATCAAGTA 2040
Qy	2041	TCCAATTTAGTGGCGTGTATTTCGGATGAATTCGCTTAGATGAAAAAGAGAGAAATTACTT 2100

Db 2041 TCCAAATTTAGTGGGTGTTTATCGGATGAATCTGCTTAGATGAAAGAGAGATTAATT 2100
Qy 2101 GAGAAAGTGAAATATGCGAAACGACTCAGTGATGAAAGAACTTACTCCAGAGATCCAAC 2160
Db 2101 GAGAAAGTGAAATATGCGAAACGACTCAGTGATGAAAGAACTTACTCCAGAGATCCAAC 2160
Qy 2161 TCCATCCATCAATAGACCAACGACTTCATATCTACTAATGAGCAATCGAATTTTCA 2220
Db 2161 TCCATCCATCAATAGACCAACGACTTCATATCTACTAATGAGCAATCGAATTTTCA 2220
Qy 2221 TCTATCCATGAACAATCTGAACATGGAATGTTGGGGAAGTGAGAACATTCACATCCAGAA 2280
Db 2221 TCTATCCATGAACAATCTGAACATGGAATGTTGGGGAAGTGAGAACATTCACATCCAGAA 2280
Qy 2281 GGAATGACGTATTTAAAGAGAAATACGTCACTACCGGGGACTTTTAAATGAGTGTAT 2340
Db 2281 GGAATGACGTATTTAAAGAGAAATACGTCACTACCGGGGACTTTTAAATGAGTGTAT 2340
Qy 2341 CCGACGTATTTATCAAAAATAGGAGAGTCCGAATTTAAAGCTTATCTCGTACCAA 2400
Db 2341 CCGACGTATTTATCAAAAATAGGAGAGTCCGAATTTAAAGCTTATCTCGTACCAA 2400
Qy 2401 TTAAGAGGTATATTGAAGATAGTCAAGATTTAGAGATATATTGATTCGTTAATGCG 2460
Db 2401 TTAAGAGGTATATTGAAGATAGTCAAGATTTAGAGATATATTGATTCGTTAATGCG 2460
Qy 2461 AAACATGAACATTTGATGTTTCCAGGTACCGAGTCCGTATGGCGCTTTTCAGTTGAAAGC 2520
Db 2461 AAACATGAACATTTGATGTTTCCAGGTACCGAGTCCGTATGGCGCTTTTCAGTTGAAAGC 2520
Qy 2521 CCAATCGAAGGTGCGGAGAACCGAATCGATCGCACACATTTGAATGGAATCCTGAT 2580
Db 2521 CCAATCGAAGGTGCGGAGAACCGAATCGATCGCACACATTTGAATGGAATCCTGAT 2580
Qy 2581 CTAGATTGTTCTCGACAGATGGAAGAAATGTCGCGCATATCCCATCAATTTCTCTTTG 2640
Db 2581 CTAGATTGTTCTCGACAGATGGAAGAAATGTCGCGCATATCCCATCAATTTCTCTTTG 2640
Qy 2641 GATATTGATATTGATGCACAGACTTGCATCAGAAATCTAGCGGTGTTGGTGTATTTCAAG 2700
Db 2641 GATATTGATATTGATGCACAGACTTGCATCAGAAATCTAGCGGTGTTGGTGTATTTCAAG 2700
Qy 2701 ATTAAGACGAGGAAGTGTATGCAAGACTAGGGAATCTGGAATTTATGAAGAGAAACCA 2760
Db 2701 ATTAAGACGAGGAAGTGTATGCAAGACTAGGGAATCTGGAATTTATGAAGAGAAACCA 2760
Qy 2761 TTATTAGAGAGCACTGTCTCGTGTGAAGAGACGAGAGAAATCGAGACAAACGT 2820
Db 2761 TTATTAGAGAGCACTGTCTCGTGTGAAGAGACGAGAGAAATCGAGACAAACGT 2820
Qy 2821 GAAAACTCAATTTGGAACAAACAGAGTATATACAGAGGCAAAAGAGCTGTGGATGCT 2880
Db 2821 GAAAACTCAATTTGGAACAAACAGAGTATATACAGAGGCAAAAGAGCTGTGGATGCT 2880
Qy 2881 TTATTTGTAGATCTCAATATATAGATTACAAGCGGATCAAAACATTTGGCATGATTCAT 2940
Db 2881 TTATTTGTAGATCTCAATATATAGATTACAAGCGGATCAAAACATTTGGCATGATTCAT 2940
Qy 2941 CGGCGAGATAAATTTGTTATCGAATTTGAGAGGCTTATCTGTCAGAAATTTATCTGTTATC 3000
Db 2941 CGGCGAGATAAATTTGTTATCGAATTTGAGAGGCTTATCTGTCAGAAATTTATCTGTTATC 3000
Qy 3001 CCGGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAAGTGCATTTATCTGCAATCTCC 3060
Db 3001 CCGGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAAGTGCATTTATCTGCAATCTCC 3060
Qy 3061 CTATACGATCGAGAAATGCTGTTAAAGATTTTAAATGATTTAATGATGATGATGCTGG 3120
Db 3061 CTATACGATCGAGAAATGCTGTTAAAGATTTTAAATGATTTAATGATGATGATGCTGG 3120
Qy 3121 AATGTAAGGCGCATGTAGATGTAACAAGAGCCATCACCGTTCTGCTTGTATCCCA 3180

Db 3121 AATGTAAGGCGCATGTAGATGTAACAAGAGCCATCACCGTTCTGCTTGTATCCCA 3180
Qy 3181 GAATGGGAAGCAGAGTGTCTCAGAGCAGTTCGCGTCTGTCGGGGCGTGCCTATATCCTC 3240
Db 3181 GAATGGGAAGCAGAGTGTCTCAGAGCAGTTCGCGTCTGTCGGGGCGTGCCTATATCCTC 3240
Qy 3241 CGTGTACACGCTGACAAAGAGGATATGAGAGGGTGTGTAACGATCCATGAATCGAG 3300
Db 3241 CGTGTACACGCTGACAAAGAGGATATGAGAGGGTGTGTAACGATCCATGAATCGAG 3300
Qy 3301 AACAAATACAGACCAACTTAAATTTAAAACTGTGAAGAGAGGAGTGTATCCAAACGGAT 3360
Db 3301 AACAAATACAGACCAACTTAAATTTAAAACTGTGAAGAGAGGAGTGTATCCAAACGGAT 3360
Qy 3361 ACAGGAACGTGTATGATTTATCTGCACACCAAGGTACAGCAGTATGTAATTCCTGTAAT 3420
Db 3361 ACAGGAACGTGTATGATTTATCTGCACACCAAGGTACAGCAGTATGTAATTCCTGTAAT 3420
Qy 3421 GCTGGATATGAGGATGTCATATGAAGTTGATCTACAGCATCTGTTAATTTACAAACCGACT 3480
Db 3421 GCTGGATATGAGGATGTCATATGAAGTTGATCTACAGCATCTGTTAATTTACAAACCGACT 3480
Qy 3481 TATCAAGAGAAACGCTATACAGATGTACGAAGAGATAATCAITTTGTAATATGACAGAGG 3540
Db 3481 TATCAAGAGAAACGCTATACAGATGTACGAAGAGATAATCAITTTGTAATATGACAGAGG 3540
Qy 3541 TATGTGAATTTATCCACATACAGCTGGTTATGACAAAGAAATTTAGAAATTTCTCCCA 3600
Db 3541 TATGTGAATTTATCCACATACAGCTGGTTATGACAAAGAAATTTAGAAATTTCTCCCA 3600
Qy 3601 GAAACCGATAAGGTATGAGATTGGAGAACCGAAGGAGTATTATTGTAGACAGC 3660
Db 3601 GAAACCGATAAGGTATGAGATTGGAGAACCGAAGGAGTATTATTGTAGACAGC 3660
Qy 3661 GTGCAATTTACTCTTATGGAGGAATAG 3687
Db 3661 ATAGATTACTCTTATGGAGGAATAG 3687

RESULT 4

AF077326 LOCUS 3684 bp DNA linear BCT 26-AUG-1998
DEFINITION Bacillus thuringiensis CryIb delta-endotoxin gene, complete cds.
ACCESSION AF077326
VERSION AF077326.1 GI:3360518
KEYWORDS
SOURCE
ORGANISM
Bacillus thuringiensis
Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE 1 (bases 1 to 3684)
AUTHORS Payne, J., Cummings, D.A., Cannon, R.J.C., Narva, K.E. and Stelman, S.J.
TITLE Bacillus thuringiensis genes encoding lepidopteran-active toxins
JOURNAL Patent: US 5723758 03-MAR-1998; Mycogen Corporation; San Diego, CA
REFERENCE 2 (bases 1 to 3684)
AUTHORS Stelman, S.J.
TITLE Direct Submission
JOURNAL Submitted (08-JUL-1998) Molecular Biology, Mycogen Corporation, 5501 Oberlin Dr., San Diego, CA 92121, USA
FEATURES
source
1. 3684
/organism="Bacillus thuringiensis"
/mol_type="genomic DNA"
/strain="PS158C2"
/db_xref="taxon:1428"
/plasmid="pMYC2383"
1. 3684
/codon_start=1
/transl_table=11
/product="CryIb delta-endotoxin"
/protein_id="AAC32850.1"
/db_xref="GI:3360519"
CDS

```

/translation="MTSNRRKNEIINALSIPAVSNHSAQNLSTDAIEDSLCAEG
NNIDPFVSAQVOTGGINAGRIILGVLPFAQVATSFYSLVGLWMPGRDVPWEIFLE
HVEOLIRFQVOTGGINAGRIILGVLPFAQVATSFYSLVGLWMPGRDVPWEIFLE
LEUDDFNAMPLEAIRNEOPLVMVYAAANLHLLLRDASLFGSEPLTQEOLORYE
RQVEKREYSDYCARWNTGSLNRLGNNAESWLRNQRDRLDLVLVDLVALPSPYDT
RVIPMTISAQUREISYITDIPGRINAPSGFASFTWNNPNSPFAIEAAYVPHLLDF
PEQITIFSVLSRNSNTQYMYWVGHRLSRTIRGSLSTHGNNTSINPVTLQFTSR
DVTPTESFAGINILLTTPVNGVPWARFNWENPLNSRLSLYITGVTGQLFDSBT
ELPQETTERNSYSYSHRLNIRLISGNTLRAPVYSWTHRSADRTNITSSDSITQPL
VKFNLNSGTSVSVSGPCTGDIIRTNVNGSVLSMGLPNNTSLOVRVVRVYAAQST
MVLAVTGGSTTDQGPSPMSANESLTSQSPFAEPFVGISASGQTAGISINNAQ
RQTFPHDKIFIPITAFAYIDLERAEAVNALFTNPNRLKTVGTVTDHIDEVSNL
VACLSDEFCDEKRELEKRYAKRLSDERLLODNFTSINKPOPIFSTNEQSNFTS
IHQSEHGWSGNSNITIQEONGKENVILPGTFNECYPTLYIKTGAEALYKAFPEW
QLSGYIEDSOLLEYLRYNKAKETLDPGTVESPLVSLVESPIGRCEPKNRCAHPFEM
NPDLDCSCDRGEKCAHSHLHPSLIDVGCIDLHENLGVWVVKIKTQEGHARLGNLEF
IEKPLLGEALSRKRAKWKDRKREKLOLETRVYKEAENDVALFVDSQYORLQND
TNTGMHAADALVHRIEALYSELVSPGVNAEIFELEGRIITALSILDARNVXNG
DFNNGLACWNVKHVDVQSHHRSVLVPEWEAEVQAVRVCPCRGYILRVYAYKEGY
GEGCVTTHETENNDELKFNCEEEVYPTDGTCDNYTAHQGTAAACNSNAGYEDAY
EVDITASNKYKPYEEETVYDVRDNHCEYDRGVNYPVPVAGVMTKELBYFPETDKV
WIEIGETEGKFIQVDSVELLMEE"

```

ORIGIN

Query Match 88.6%; Score 3266.6; DB 1; Length 3684;
 Best Local Similarity 93.4%; Pred. No. 0;
 Matches 3449; Conservative 0; Mismatches 229; Indels 15; Gaps 3;

QY	1	TTGACTTCAATAGGAAATGAGAAATGAAATTAATAATGCTTTATCGATTCAGCTGTA	60
DB	1	TTGACTTCAATAGGAAATGAGAAATGAAATTAATAATGCTTTATCGATTCAGCTGTA	60
QY	61	TCGAATCATTCACACAAATGATCTATCACCAGATGCTGTAATGAGGATCTTTGTGT	120
DB	61	TCGAATCATTCACACAAATGATCTATCACCAGATGCTGTAATGAGGATAGCTTGTGT	120
QY	121	ATAGCCGAGGGGAATAATATCAATCCACTTTGTTAGCGCATCAACAGTCCAAACGGGTATT	180
DB	121	ATAGCCGAGGGGAATAATATCGATCCACTTTGTTAGCGCATCAACAGTCCAAACGGGTATT	180
QY	181	AACATAGCTGTAGAAATCTAGGTGATATAGGCGTACGGTTGCTGACAAATAGCTAGT	240
DB	181	AACATAGCTGTAGAAATCTAGGTGATATAGGCGTACGGTTGCTGACAAATAGCTAGT	240
QY	241	TTTTATAGTTTCTTGTGTGTGTAATATGCGCCGCGGAGAGATCAGTGGGAAATTTTC	300
DB	241	TTTTATAGTTTCTTGTGTGTGTAATATGCGCCGCGGAGAGATCCTTGGGAAATTTTC	300
QY	301	CTAGAACATGTGGAACAACTTTATAATCAACAAATAACAGAAATGCTAGGAATACGGCA	360
DB	301	CTAGAACATGTGGAACAACTTTATAAGAACAAACAGTAACAGAAATACGGGATACGGCT	360
QY	361	CTTGCTCGATTCAGGATTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT	420
DB	361	CTTGCTCGATTCAGGATTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT	420
QY	421	TGGCTAGAAAACCGTGATGATGCAAGAACGAGAGTGTCTTTATACCCCAATATATAGCC	480
DB	421	TGGCTAGAAAACCGTGATGATGCAAGAACGAGAGTGTCTTTATACCCCAATATATAGCC	480
QY	481	TTAGAACTTGATTTCTTAATGCGATCGCGCTTTTCGCAATTAAGAACCAAGAGTTTCCA	540
DB	481	TTAGAACTTGATTTCTTAATGCGATCGCGCTTTTCGCAATTAAGAACCAAGAGTTTCCA	540
QY	541	TTATTAATGATGATGCTCAAGCTGCAAAATTTACCTATTTATTTAGAGATGCTCTCT	600
DB	541	TTATTAATGATGATGCTCAAGCTGCAAAATTTACCTATTTATTTAGAGATGCTCTCT	600
QY	601	CTTTTTCGTAGTGAATTTGGGCTTTACATCGCAGGAAATTCAACTTATTTATGAGCGCAA	660
DB	601	CTTTTTCGTAGTGAATTTGGGCTTTACATCCNAGAAATTCACGTTATTTATGAGCGCAA	660
QY	661	GTGGAACAAACGAGAGATTTATTCGACTATTCGAGTGGTATTAATAACAGGTCTAAAT	720

DB	661	GTGGAACAAACGAGAGATTTATTCGACTATTCGAGTGGTATTAATAACAGGTCTAAAT	720
QY	721	AGCTTGAGGGGCAAAATGCCCAAGCTTGGTGGTGTATTAATCAATTCCTGAGAGATCTA	780
DB	721	AATTTGAGAGGGGCAAAATGCTGAAAGTTGGTGGGATTAATCAATTCCTGAGAGATCTA	780
QY	781	ACGTTAGGGGTATTAGATCTAGTGGCACTATTCGCCAGCTATGACACTCCGACTTATCCA	840
DB	781	ACGTTAGGAGTATTAGATCTAGTGGCACTATTCGCCAGCTATGACAGCGCTGTTTATCCA	840
QY	841	ATAAATACGAGTCTCAGTTTAAACAAGGAAAGTTTATACAGACGCAATTTGAGCAACAGGG	900
DB	841	ATGAATACCACTGCTCAATTTAAACAAGAAATTTATACAGATCAATTTGGGAGAACAAAT	900
QY	901	GTAAT-----ATGCAAGTATGAATTTGGTATAATAATCACTTCCGTTTCCGCT	954
DB	901	GCACCTTCAGGATTTGCAAGTACGATTTGGTGTAAATAATGACCATCGTTTTCTGCC	960
QY	955	ATAGAGACTCGCGTTATCCGAAGCCGCACTACTTGAATTTCTAGAAACAATTTACAAT	1014
DB	961	ATAGAGCTGCCGTTATAGGCCCTCCGCACTACTTGAATTTCCAGAACAGCTTACAAT	1020
QY	1015	TTTAGCATTTCATCAGATGAGTGTCTATAGGCAATATGACTTACTGCGGGGGGACACA	1074
DB	1021	TTTAGCGCTATTAGTTCGATGAGTAAATCTCAATATATGAAATTTACTGGGTGGGACATAGA	1080
QY	1075	ATTCAATCTCGGCCAATAGGAGCGGATTAATAACCTCAACGATGCTGCTACCAATACT	1134
DB	1081	CTTGAATCGGCAACAAATAAGGGGTCTTAATGATCTCGACACACGGAATATCAATACT	1140
QY	1135	TCATTAAATCTCTGAAGATTTATCACTTCTCTCGAGACGTAATTTGGACTGAATCATAT	1194
DB	1141	TCATTAAATCTCTGAATTTACAGTTTCACTCTCGAGACGTTTATAGAACAGAAATCATTT	1200
QY	1195	GACGAGTGTCTTATGAGGGAATTTACCTTGAACCTTATTCATGCTGCTCCTACTGTTAGA	1254
DB	1201	GCAGGGATAAATAT-----ACTTCTAACTACTCTCTGTAATGGAGTACCTTGGGCTAGA	1254
QY	1255	TTTAAATTTTAGGAAACCTCAGAAATCTTTTGAAGAGGTACTCTAACTATATAGTCAACCC	1314
DB	1255	TTTAAATTTTAGGAAATCTCTGAAATCTCTT---AGAGGTAGCCTTCTTATATATAGG	1311
QY	1315	TATAGTCTACCTGGGCTTCAATTTAAAGATTTCAAGAAATTCAGAAATTCACACAGAAACA	1374
DB	1312	TATAGTGGAGTGGGACACAACTTATTTGATTCAGAAATTCAGAAATTCACACAGAAACA	1371
QY	1375	GAACGACCAATTTATGAATCATATAGTCTATAGTCTTATCTCAGATGAGGCTATTTACAA	1434
DB	1372	GAACGACCAATTTATGAATCTTACAGTCTATAGTCTTATCTAATAATAAGACTTAATCAGGA	1431
QY	1435	TCTAGGCTGATCTACAGTATATTTCTTGGACGACCGTAGTGCAGATCGTACAAATACC	1494
DB	1432	AACACTTGGAGACACAGTATATTTCTTGGACGACCGTAGTGCAGATCGTACAAATACC	1491
QY	1495	ATTAGTTTCAGATAGCATAACACAAATACCAATTTGGTAAATTCATTCAACCTTAAATCAGGT	1554
DB	1492	ATTAGTTTCAGATAGCATAACACAAATACCAATTTGGTAAATTCATTCAACCTTAAATCAGGT	1551
QY	1555	ACCTCTGTAGTCACTGCGCCAGGATTTACAGGAGGGGATATAATTCGAACTAAGCTTAAT	1614
DB	1552	ACCTCTGTAGTCACTGCGCCAGGATTTACAGGAGGGGATATAATTCGAACTAAGCTTAAT	1611
QY	1615	GGTAGTGTATTAAGTATGGTCTTAATTTTAAATAATCAATCAATACAGCGGTATCGCGT	1674
DB	1612	GGTAGTGTATTAAGTATGGTCTTAATTTTAAATAATCAATCAATACAGCGGTATCGCGT	1671
QY	1675	AGAGTTTGGTATGCTCTTCTCAACAAATGGTCTCGAGGGTAACTGTGCGAGGGAGTACT	1734
DB	1672	AGAGTTTGGTATGCTCTTCTCAACAAATGGTCTCGAGGGTAACTGTGCGAGGGAGTACT	1731
QY	1735	ACTTTTGTATCAAGGATTTCCCTAGTACTATGAGTGCAAATGAGTCTTTGACATCTCAATCA	1794

1732	ACTTTTGATCAAGGATTCCCTAGTACTATGAGTGCAAATGAGCTTTTGGACATCTCAATCA	1791
1795	TTTAGATTTGCAGAAATTTCCCTGTAGGTATTTAGTGCATCTGGCAGTCAAACTGCTGGAATA	1854
1792	TTTAGATTTGCAGAAATTTCCCTGTAGGTATTTAGTGCATCTGGCAGTCAAACTGCTGGAATA	1851
1855	AGTATAAGTAAATAATGCAGGTAGACAAACGTTTCACCTTTGATATAAAATTGAATTCATTTCCA	1914
1852	AGTATAAGTAAATAATGCAGGTAGACAAACGTTTCACCTTTGATATAAAATTGAATTCATTTCCA	1911
1915	ATTACTGCAACCTTCGAAGCAGATAACGATTTAGAAAGGCGCAAGAGGCGGTGAATGCT	1974
1912	ATTACTGCAACCTTCGAAGCAGATAATGATTTAGAAAGGCGCAAGAGGCGGTGAATGCT	1971
1975	CTGTTTACTAATACGAATCCAAAGAAGATTGAAAAACAGATGTGACAGATTAATCATATTGAT	2034
1972	CTGTTTACTAATACGAATCCAAAGAAGGTTGAAAAACAGGTGTGACAGATTAATCATATTGAT	2031
2035	CAAGTATCCAAATTTAGTGGCGTGTATTATCCGATCGAATTTCTGCTTTAGATGAAAAGAGAGAA	2094
2032	GAAGTATCCAAATTTAGTGGCGTGTATTATCCGATCGAATTTCTGCTTTGAGTAAAAAGAGAGAA	2091
2095	TTACTTCGAGAAAGTGAATAATGCGAAACGACTCAGTGCATGAAAGAACTTTACTCCAGAT	2154
2092	TTACTTCGAGAAAGTGAATAATGCGAAACGACTCAGTGCATGAAAGAACTTTACTCCAGAT	2151
2155	CCAACTTTCACATCCATCAATTAAGCAACACAGACTTCATATCTACTAATGAGCAATCGAAT	2214
2152	CCAACTTTCACATCCATCAATTAAGCAACACAGACTTCATATCTACTAATGAGCAATCGAAT	2211
2215	TTACATCTTATCCATGAACAATCTGAAACGATGTTGGGGAAGTGAGAACTTTACAATC	2274
2212	TTACATCTTATCCATGAACAATCTGAAACGATGTTGGGGAAGTGAGAACTTTACAATC	2271
2275	CAGCAAGAAATGACGTTATTTAAAGAGAAATTAAGTCAACATACCGGGGACTTTTAATGAG	2334
2272	CAGCAAGAAATGACGTTATTTAAAGAGAAATTAAGTCAACATACCGGGGACTTTTAATGAG	2331
2335	TGTTATCCGACGTTATTTATATCAAAAAATAGGAGAGTCGGAATTTAAAGCTTATATCTCGC	2394
2332	TGTTATCCGACGTTATTTATATCAAAAAATAGGAGAGTCGGAATTTAAAGCTTATATCTCGC	2391
2395	TACCAATTAAGAGGTATATTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT	2454
2392	TACCAATTAAGTGGCTATATTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT	2451
2455	AATCGGAAACATGAAACATTTGGATGTTCCAGGTACCGAGTCCGTATGCCCGCTTTCAGTT	2514
2452	AATCGGAAACATGAAACATTTGGATGTTTCCAGGTACCGAGTCCGTATGCCCGCTTTCAGTT	2511
2515	GAAAGCCAAATCGGAAGGTGCGGAGAACCGAATCGATCGCACACACATTTTGAATGGAAT	2574
2512	GAAAGCCAAATCGGAAGGTGCGGAGAACCGAATCGATCGCACACACATTTTGAATGGAAT	2571
2575	CCTGATCTAGATTGTTCTCTCAGAGATGGAGAAAAATGTGCGCATCAATCCCATCATTTTC	2634
2572	CCTGATCTAGATTGTTCTCTCAGAGATGGAGAAAAATGTGCGCATCAATCCCATCATTTTC	2631
2635	TCCTTTGATATTTGATTTGGATGCACAGCTTGCAATGAGAAATCTAGGCGGTGTGGTGGTA	2694
2632	TCCTTTGATATTTGATTTGGATGCATAGACTTTGCAAGAACCTTAGGCGGTGTGGTGGTA	2691
2695	TTCAAGATTTAAGACGAGGAGGTTCATGACACTAGGAAATCTGGAATTTTATTGAAGAG	2754
2692	TTCAAGATTTAAGACGAGGAGGTTCATGACACTAGGAAATCTGGAATTTTATTGAAGAG	2751
2755	AAACCAATTTATTAGGAGAGCACTCTCTCGTGTGAAGAGACAGAGAAAAAATGGAGAGAC	2814
2752	AAACCAATTTATTAGGAGAGCACTCTCTCGTGTGAAGAGACAGAGAAAAAATGGAGAGAC	2811
2815	AAACGTGAAAAACTCAATTTGGAAAAACAAACAGATATATACAGAGGCAAAAGAGCTGTG	2874
2812	AAACGTGAAAAACTCAATTTGGAAAAACAAACAGATATATACAGAGGCAAAAGAGCTGTG	2871

QY	2875	GATGCTTTATTTGTPAGATTCTCAATATAATAGATTACAAGCGGATACAAAACATTGGCATG	2934
Db	2872	GATGCTTTATTTGTPAGATTCTCAATATGATAGATTACAAGCGGATACAAAACATTGGCATG	2931
QY	2935	ATTTCATCGGCGAGATAAATCTGTTTCATCGAATTCGAGAGCGCTTATCTGTGAGAAATTAATCT	2994
Db	2932	ATTTCATCGGCGAGATAAATCTGTTTCATCGAATTCGAGAGCGCTTATCTTTTTCAGAAATTAATCT	2991
QY	2995	GTTATCCCGGGTGTAATTCGGGAAATTTTGAAGAAATTTAGAAAGTTCGATTAATAATGGAATTAATCT	3054
Db	2992	GTTATCCAGGTGTAATTCGGGAAATTTTGAAGAAATTTAGAAAGTTCGATTAATAATGGAATTAATCT	3051
QY	3055	ATTCCTCCTATACGATGCGAGAAATGTCGTTTAAAAATCGTGAATTTTAATAATGGAATTAATCT	3114
Db	3052	ATTCCTCCTATACGATGCGAGAAATGTCGTTTAAAAATCGTGAATTTTAATAATGGAATTAATCT	3111
QY	3115	TGCTGGAAATGTAAGAAGGCGATGATGTTACACAGAGCCCATCACCGTTCTGTCTCTTGT	3174
Db	3112	TGCTGGAAATGTAAGAAGGCGATGATGTTACACAGAGCCCATCACCGTTCTGTCTCTTGT	3171
QY	3175	ATCCCAATGCGGAGCAGAAAGTGTCAACAGCAGTTCCGCTCTGTCCGGGCGTGGCTAT	3234
Db	3172	ATCCCAATGCGGAGCAGAAAGTGTCAACAGCAGTTCCGCTCTGTCCGGGCGTGGCTAT	3231
QY	3235	ATTCCTCGTGTCAAGCGGTCAAAAGAGGGATATGAGAGGGTGTGTGAACGATCCATGAA	3294
Db	3232	ATTCCTCGTGTCAAGCGGTCAAAAGAGGGATATGAGAGGGTGTGTGAACGATCCATGAA	3291
QY	3295	ATCGAGAAACATATACAGACGAACTTAAATTTTAAAACTGTGAAGAGGAGTGTATCCCA	3354
Db	3292	ATCGAGAAACATATACAGACGAACTTAAATTTTAAAACTGTGAAGAGGAGTGTATCCCA	3351
QY	3355	ACGATACAGAAAGCTGTAATGATTATCTACACCAAGTGTGACAGCAGTATGTAATTC	3414
Db	3352	ACGATACAGAAAGCTGTAATGATTATCTACACCAAGTGTGACAGCAGTATGTAATTC	3411
QY	3415	CGTAATGCTGGATATGAGATGTCATATGAAGTTGATACTACAGCATCTGTTAATTAACAA	3474
Db	3412	CGTAATGCTGGATATGAGATGTCATATGAAGTTGATACTACAGCATCTGTTAATTAACAA	3471
QY	3475	CCGACTTATGAAGAGAAACGTTATACAGATGTACGAGAGATATCATCTGTGAATATGAC	3534
Db	3472	CCGACTTATGAAGAGAAACGTTATACAGATGTACGAGAGATATCATCTGTGAATATGAC	3531
QY	3535	AGAGGATGTGAATATCCACCACTACAGCTGTTTATATGACAAAGAAATAGAAATAC	3594
Db	3532	AGAGGATGTGAATATCCACCACTACAGCTGTTTATATGACAAAGAAATAGAAATAC	3591
QY	3595	TTCCAGAAACCCGATAAGTATGGAATGAGATTGGAGAAACCGAAGGAAAGTTTATTTGTA	3654
Db	3592	TTCCAGAAACCCGATAAGTATGGAATGAGATTGGAGAAACCGAAGGAAAGTTTATTTGTA	3651
QY	3655	GACAGCGTGAATTAATCTCTTATGAGGAAATAG	3697
Db	3652	GACAGCGTGAATTAATCTCTTATGAGGAAATAG	3684
RESULT 5			
LOCUS			
190319			
DEFINITION			
Sequence 7 from patent US 5723758.			
ACCESSION			
190319			
VERSION			
190319.1			
KEYWORDS			
GI:3410259			
SOURCE			
Unknown.			
ORGANISM			
Unknown.			
REFERENCE			
1 (bases 1 to 3684)			
AUTHORS			
Payne,J., Cummings,D.A., Cannon,R.J.C., Narva,K.E. and Stelman,S.			
TITLE			
Bacillus thuringiensis genes encoding lepidopteran-active toxins			
JOURNAL			
Patent: US 5723758-A 7 03-MAR-1998;			
FEATURES			
Location/Qualifiers			
linear			
PAT 10-AUG-1998			

	source	1..3684 /organism= "unknown" /mol_type= "unassigned DNA "
	ORIGIN	
	Query Match	88.6%; Score 3266.6; DB 6; Length 3684;
	Best Local Similarity	93.4%; Pred. No. 0;
	Matches 3449; Conservative	0; Mismatches 229; Indels 15; Gaps 3
Qy	1	TTGACTTCAAATAGGAAAAATGAGAATGAAATTATAAATGCTTTATCGATTCCAGCTGTGA 60
Db	1	TTGACTTCAAATAGGAAAAATGAGAATGAAATTATAAATGCTTTATCGATTCCAGCTGTGA 60
Qy	61	TGCAATCATNTCCA CAAAATGGATCTATACA CAGATGCTCGTATTGAGAGATTCTTTTGCT 120
Db	61	TGCAATCATNTCCGACAAATGAAATCTATCAACCGATGCTCGTATTGAGGATAGCTTGCT 120
Qy	121	ATAGCCGAGGGGAATAATCAATCCACTGTTGTAGGCATCAACAGTCACCAACGGGTATT 180
Db	121	ATAGCCGAGGGGAACAATCGATCCATVTGTTAGGCATCAACAGTCACCAACGGGTATT 180
Qy	181	AACATAGCTGTTAGAAATACTAGTGTAATTAGCGGTACCCTTTGCTGGACAAATAGTAGT 240
Db	181	AACATAGCTGTTAGAAATACTAGTGTAATTAGCGGTACCCTTTGCTGGACAAATAGTAGT 240
Qy	241	TTTTATAGTTTTCTTTGTTGGTGAATTATGGCCCCGCCGCGCAGAGATCAGTGGGAAATTTTC 300
Db	241	TTTTATAGTTTTCTTTGTTGGTGAATTATGGCCCCGCCGCGCAGAGATCCTTTGGGAAATTTTC 300
Qy	301	CTAGAACATGTCGAACAACCTTTATAATCAACAATAACAGAAAAATGCTTAGGAATACGGCA 360
Db	301	CTAGAACATGTCGAACAACCTTTATAAGACAA CAAGTAA CAGAAAAATATAGGATACGGCT 360
Qy	361	CTTGCTCGATTA CAAGGTTTATGGAGATTCTTTAGAGCCTATCAACAGTCACCTTGAAGAT 420
Db	361	CTTGCTCGATTA CAAGGTTTATAGGAAATTCCTTTAGAGCCTATCAACAGTCACCTTGAAGAT 420
Qy	421	TGGCTAGAAAACCGTGATGATGCAAGACGAGAAGTGTTCTTTATACCAATATATATAGCC 480
Db	421	TGGCTAGAAAACCGTGATGATGCAAGAA CGAGAAGTGTTCTTTATACCCAATATATATAGCC 480
Qy	481	TTAGAACTTGATTTTCTTTAATGCGATGCGCTTTTCGCAATTAGAAAA CCAAGAAAGTTCCA 540
Db	481	TTAGAACTTGATTTTCTTTAATGCGATGCGCTTTTCGCAATTAGAAA CCAAGAGATTTCCA 540
Qy	541	TTATTAAATGGTATATGCTCAAGTCGCAAAATTTACACCTATTATTATTGAGAGATGCCTCT 600
Db	541	TTATTAAATGGTATATGCTCAAGTCGCAAAATTTACACCTATTATTATTGAGAGATGCCTCT 600
Qy	601	CTTTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACGTTATTATGAGCGCCAA 660
Db	601	CTTTTTTGGTAGTGAATTTGGGCTTACATCCCAGAAATTTCAACGTTATTATGAGCGCCAA 660
Qy	661	GTCGAA CAAACACGAGAGATTATTCGACTATTTCGCTAGAAATGGGTATAAATACAGCTCTAAAT 720
Db	661	GTCGAA CAAACACGAGAGATATTCTGATTATTTCGCGCAGATGGTATPATACGGTTTAAAT 720
Qy	721	AGCTTGAGAGGGA CAAATGCCCGCAAGTTTGGGTGCGTTATAATCAATTCGTTAGAGATCTTA 780
Db	721	AAATTTGAGAGGGA CAAATGCTGAAAGTTGGTTGCGATATAATCAATTCGTTAGAGATCTTA 780
Qy	781	ACGTTAGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGACACTCGCACTTATCCA 840
Db	781	ACGCTTAGAGGTATTAGATCTCTAGTGGCACTATTTCCTCAAGCTATGACACCGCTTTATCCA 840
Qy	841	ATAAATACGAGTGTCTCAGTTAAACAGGGAAGTTTATACAGACGCAATTTGGACCAACAGGG 900
Db	841	ATGAATACCACTGCTCAATTAACAGAGAAATTTATACAGATCCCAATTCGGGAGAACAAAT 900
Qy	901	GTAAT-----ATGGCAAGTATGAATTTGGTATAATAATAATATGCACTTTCGTTTTCCGCT 954
Db	901	GCACCTTCAGGATTTGCAAGTACGAATTTGGTTTAAATAATGACCACTCGTTTTCTGCC 960

Qy	955	ATAGAGACTCGGGTATTCGAAAGCCGCATCTACTACTGATTTTCTAGAAACAACTTCAAAATT	101
Db	961	ATAGAGGCTCGCGTTATTTAGGCGCTCCGCATCTACTTGAATTTTCCAGAAACAGCTTCAAAATT	102
Qy	1015	TTTAGCAGCTTCATCACGATCGAGTGCTACTAGGCATATGACTTACTTGGCGGGGCGCACACA	1074
Db	1021	TTTACGCGTATTAAGTCGATGGAGTAATATCTCAATATATGAAATTTACTGGGTGGGACATAGA	1080
Qy	1075	ATTCAATCTTCGGCCAATPAGGAGCGGATTAAATACCTCAACGCGATGGGTCTACCAATACT	1134
Db	1081	CTTGAATCGGAACAATAAGGGGTCATTAAAGTACCTCGACACACGCGAATACCAATACT	1140
Qy	1135	TCATTAATCCTGTAGATTATCATTTCTCTCGACGCTATATTGGACTGAATCATAT	1194
Db	1141	TCATTAATCCTGTAAATTTACAGTTTACATCTCGAGACGTTTTATAGAACAGAAATCATTTT	1200
Qy	1195	GCAGGAGTCTTCTATGGGGAAATTTACCTTGAACCTTATCATGGTGTCCCTACTGTTTAGA	1254
Db	1201	GCAGGGATAAATAT-----ACTTCTAACTACTCTCTGTGAATGGAGTACCTTGGGCTAGA	1254
Qy	1255	TTTAAATTTTAGGAACCTTCAGAAATACTTTTGAAGAGGTACTGCTAACTATAGTCAACCC	1314
Db	1255	TTTAAATTTAGGAATCCCTGAAATCTCTT--AGAGTAGCCCTTCTCTATCTATAGGG	1311
Qy	1315	TATGAGTCACTCGGGCTTCAATTTAAAGATTTCAGAACTGAAATTAACAACAGAAACAACA	1374
Db	1312	TATACTGGAGTGGGACACAACCTATTTTGATTCAGAAATGNAATTAGCACCAGAAACAACA	1371
Qy	1375	GAACGACCAATTTATGAATCATATAGTATAGTATCTCACATAGGGCTCATTTCCACA	1434
Db	1372	GAACGACCAATTTATGAATCTTACAGTCATAGATTAATCTAAATATAGACTTAATCAGGA	1431
Qy	1435	TCTAGGCTGCATGTACCAGTATATCTTGAGCACCGCTAGTCAGATCGTACAAATACC	1494
Db	1432	AACACTTTGAGACGACCAAGTATATCTTGAGACGACGCTAGTCAGATCGTACAAATACC	1491
Qy	1495	ATTAGTTTCAGATAGCATAAACAAATAACCATCGGTAAAAATCATTCAACCTTAATCAGGT	1554
Db	1492	ATTAGTTTCAGATAGCATAAACAAATACCATCGGTAAAAATCATTCAACCTTAATCAGGT	1551
Qy	1555	ACCTCTGTAGTCACTGGCCAGAGATTTACAGGAGGGGATATAATCCGAACCTAACGTTAAT	1614
Db	1552	ACCTCTGTAGTCACTGGCCAGAGATTTACAGGAGGGGATATAATCCGAACCTAACGTTAAT	1611
Qy	1615	GGTAGTGTACTAAGTATGGGTCTTAATTTTAAATAACATCATTTACAGCGGTATCGCGTG	1674
Db	1612	GGTAGTGTACTAAGTATGGGTCTTAATTTTAAATAACATCATTTACAGCGGTATCGCGTG	1671
Qy	1675	AGAGTTTCGTTATCTGCTTCTCAAAACAATGGTCTCGAGGGTAACTGTGCGAGGGAGTACT	1734
Db	1672	AGAGTTTCGTTATCTGCTTCTCAAAACAATGGTCTCGAGGGTAACTGTGCGAGGGAGTACT	1731
Qy	1735	ACTTTTGTATCAAGGATTCCTAGTACTATAGTGTGCAAAATGAGTCTTTTGAATCTCAATCA	1794
Db	1732	ACTTTTGTATCAAGGATTCCTAGTACTATAGTGTGCAAAATGAGTCTTTTGAATCTCAATCA	1791
Qy	1795	TTTTAGATTTTCAGAAATTTCTGTAGTATTTAGTGCATCTGGCGAGTCAAACTGCTGGAATA	1854
Db	1792	TTTTAGATTTTCAGAAATTTCTGTAGTATTTAGTGCATCTGGCGAGTCAAACTGCTGGAATA	1851
Qy	1855	AGTATAAGTATAATGCAAGTACAAAAACGTTTTCACTTTGATAAAAATGAAATTCATTCCA	1914
Db	1852	AGTATAAGTATAATGCAAGTACAAAAACGTTTTCACTTTGATAAAAATGAAATTCATTCCA	1911
Qy	1915	ATTACTGCAACCTTTGGAAGCAGAAATACGATTTAGAAAGGGCGCAGAGCGGTGAATGCT	1974
Db	1912	ATTACTGCAACCTTTGGAAGCAGAAATACGATTTAGAAAGGGCGCAGAGCGGTGAATGCT	1971
Qy	1975	CTGTTTTACTAATACGAATCCAAGAGATTGAAACACAGATGTGACAGATTATCATATTGAT	2034
Db	1972	CTGTTTTACTAATACGAATCCAAGAGATTGAAACACAGATGTGACAGATTATCATATTGAT	2031
Qy	2035	CAAGTATCCAATTTTAGTGGCGTGTATTATCGGATGAAATTTCTGCTTAGATGAAAGAGAGAA	2094

```
Db 2032 GAAGTATCCAAATTAGTGGCGTGTATCGATGAAATCTGCTTGGATGAAGAAGAGAGAA 2091
Qy 2095 TTACTTCAGAAAGTGAATATGCGAAACGACTCAGTGCATGAAAGAAACTTACTTCCAAGAT 2154
Db 2092 TTACTTCAGAAAGTGAATATGCGAAACGACTCAGTGCATGAAAGAAACTTACTTCCAAGAT 2151
Qy 2155 CCAAACTTCATCATCATCAATTAAGCAACACAGACTTCAATATCTACTATGAGCAATCGAAT 2214
Db 2152 CCAAACTTCATCATCATCAATTAAGCAACACAGACTTCAATATCTACTATGAGCAATCGAAT 2211
Qy 2215 TTCACTCTATCCATGAACAACTGAACTGATGAGTGGGGAAGTGAAGACATTTACATCAATC 2274
Db 2212 TTCACTCTATCCATGAACAACTGAACTGATGAGTGGGGAAGTGAAGACATTTACATCAATC 2271
Qy 2275 CAGGAAGGAAATGACGATTTTAAAGAGAAATTAACGTCACTACCGGGGAGCTTTTAAATGAG 2334
Db 2272 CAGGAAGGAAATGACGATTTTAAAGAGAAATTAACGTCACTACCGGGGAGCTTTTAAATGAG 2331
Qy 2335 TGTATCCGAGGTATTTATATCAAAAAATAGGAGAGTCCGAAATTAAGACTTATACCTCGC 2394
Db 2332 TGTATCCGAGGTATTTATATCAAAAAATAGGAGAGTCCGAAATTTAAAGACTTATACCTCGC 2391
Qy 2395 TACCAATTAAGAGGTATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT 2454
Db 2392 TACCAATTAAGGTGCTATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT 2451
Qy 2455 AATCGGAAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTTCAGTT 2514
Db 2452 AATCGGAAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTTCAGTT 2511
Qy 2515 GAAAGCCCAATCGGAAGTGGGAGAACCGAATCGATCGGACCAACATTTTGAATGGAAT 2574
Db 2512 GAAAGCCCAATCGGAAGTGGGAGAACCGAATCGATCGGACCAACATTTTGAATGGAAT 2571
Qy 2575 CCTGATCTAGATTTCTCTCGAGAGATGGAGAAAATGTCGCGATCATTTCCCATCATTC 2634
Db 2572 CCTGATCTAGATTTCTCTCGAGAGATGGAGAAAATGTCGCGATCATTTCCCATCATTC 2631
Qy 2635 TCTTTGGATATTTGATTTGGATGCACAGACTTGCATGAGAACTTAGCGGTGGTGGTA 2694
Db 2632 TCTTTGGATATTTGATTTGGATGCATAGACTTGCATGAGAACTTAGCGGTGGTGGTA 2691
Qy 2695 TTCAAGATTAAGACCGCAGGAAGTCTATGCAAGACTAGGGAATCTGGAATTTATTTGAAGAG 2754
Db 2692 TTCAAGATTAAGACCGCAGGAAGTCTATGCAAGACTAGGGAATCTGGAATTTATTTGAAGAG 2751
Qy 2755 AAACCATTTATAGGAGAGCACTCTCTGCTGTAAGAGAGCAGAGAAAAAATGGAGAGAC 2814
Db 2752 AAACCATTTATAGGAGAGCACTCTCTGCTGTAAGAGAGCAGAGAAAAAATGGAGAGAC 2811
Qy 2815 AAACGTGAAAAAATCAAAATTTGAAACAAACGAGTATATACAGAGGCAAAAGAGCTGTG 2874
Db 2812 AAACGTGAAAAAATCAAAATTTGAAACAAACGAGTATATACAGAGGCAAAAGAGCTGTG 2871
Qy 2875 GATGCTTTTATTTGATGATTTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATG 2934
Db 2872 GATGCTTTTATTTGATGATTTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATG 2931
Qy 2935 ATTCAATCGGAGAGATAAATTTGTTTCATCGAATTCGAGAGGCTTATCTGTCAGAAATTAATCT 2994
Db 2932 ATTCAATCGGAGAGATAAATTTGTTTCATCGAATTCGAGAGGCGTATCTTTTCAAGAAATTAATCT 2991
Qy 2995 GTTATCCGGGTGTAATGCGGAATTTTGAAGAAATTTGAAGGTTCGATTTATCACTGCA 3054
Db 2992 GTTATCCGGGTGTAATGCGGAATTTTGAAGAAATTTTGAAGAAATTTGAAGGTTCGATTTATCACTGCA 3051
Qy 3055 ATCTCCCTATACGATGCGAGAAATGTCGTTAAAAAATGTCGATTTTAAATAATGGAATTAGCA 3114
Db 3052 ATCTCCCTATACGATGCGAGAAATGTCGTTAAAAAATGTCGATTTTAAATAATGGAATTAGCA 3111
Qy 3115 TGCTGGAATGTAAGAGGCAATGATAGTGTACAAACAGAGCCATCACCGTTCCTGCTTGT 3174
```

```
3112 TGCTGGAATGTAAGAGGCAATGATAGTGTACAAACAGAGCCATCACCGTTCCTGCTTGT 3171
Qy 3175 ATCCAGATGGGAGAGAGAGAGTGTCAAGAGAGTTCGCGTCTGTCCGGGCGTGGCTAT 3234
Db 3172 ATCCAGATGGGAGAGAGAGTGTCAAGAGAGTTCGCGTCTGTCCGGGCGTGGCTAT 3231
Qy 3235 ATCTCCGCTGCACAGCGTACAAAGAGGGATATGGAGAGGTTGTGTAAACGATCCATGAA 3294
Db 3232 ATCTCCGCTGCACAGCGTACAAAGAGGGATATGGAGAGGTTGTGTAAACGATCCATGAA 3291
Qy 3295 ATCGAGAACAAATACAGACGAACTAAAATTTAAAAAATCTGTGAAGAGGAGTGTATCCA 3354
Db 3292 ATCGAGAACAAATACAGACGAACTAAAATTTAAAAAATCTGTGAAGAGGAGTGTATCCA 3351
Qy 3355 ACGGATACAGAAACGTTGTAATGATTAATCTGCAACCAAGGTACAGCAGCATGTAATTC 3414
Db 3352 ACGGATACAGAAACGTTGTAATGATTAATCTGCAACCAAGGTACAGCAGCATGTAATTC 3411
Qy 3415 CGTAATGCTCGATATGAGGATGCATATGAAGTTGATACAGCATCTGTTAATTACAA 3474
Db 3412 CGTAATGCTCGATATGAGGATGCATATGAAGTTGATACAGCATCTGTTAATTACAA 3471
Qy 3475 CCGACTTTATCAAGAAAGAAACGTATACAGATGTACGAGAGATAATCATTTGTGAATATGAC 3534
Db 3472 CCGACTTTATCAAGAAAGAAACGTATACAGATGTACGAGAGATAATCATTTGTGAATATGAC 3531
Qy 3535 AGAGGTATGTGAATTTATCCACCATCCAGCTGGTTATATGACAAAGAAATTTAGATATAC 3594
Db 3532 AGAGGTATGTGAATTTATCCACCATCCAGCTGGTTATATGACAAAGAAATTTAGATATAC 3591
Qy 3595 TTCCAGAAACCGATAGGTATGAGTTGAGATTGGAGAAACCGAAGGAGTTTATTGTA 3654
Db 3592 TTCCAGAAACCGATAGGTATGAGTTGAGATTGGAGAAACCGAAGGAGTTTATTGTA 3651
Qy 3655 GACAGCGTGAATTTACTCTTTATGGAGGAATAG 3687
Db 3652 GACAGCGTGAATTTACTCTTTATGGAGGAATAG 3684
```

RESULT 6
AR359400 LOCUS 3684 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 62 from patent US 6593293.
ACCESSION AR359400
VERSION AR359400.1 GI:33765710
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3684)
AUTHORS Baum,J.A., Chu,C.-R., Donovan,W.P., Gilmer,A.J. and Rupar,M.J.
TITLE Lepidopteran-active Bacillus thuringiensis .delta.-endotoxin
compositions and methods of use
JOURNAL Patent: US 6593293-A 62-15-JUL-2003;
FEATURES
source location/Qualifiers
1. 3684
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 88.4%; Score 3260.2; DB 6; Length 3684;
Best Local Similarity 93.3%; Pred. No. 0;
Matches 3445; Conservative 0; Mismatches 233; Indels 15; Gaps 3;

Qy 1 TTGACTTCAATAGGAAAAATGAGATGAATTAATTAATGCTTTATCGATTCCAGCTGTA 60
Db 1 TTGACTTCAATAGGAAAAATGAGATGAATTAATTAATGCTTTATCGATTCCAGCTGTA 60
Qy 61 TCGAATCATTTCCACACAAATGGATCTATCCAGATGCTCGTATTTGAGGATTTCTTTGTT 120
Db 61 TCGAATCATTTCCGACACAAATGAATCTATCAACCGATGCTCGTATTTGAGGATAGCTTGT 120
Qy 121 ATAGCCGAGGGGAATATATCAATCCACTGTTTGTAGCGCATCAACAGTCCCAACCGGTATT 180

121 ATAGCCGAGGGGAA CAATATATCGATCCATTTGTTAGCGCATCAACAGTCCAAACGGGTATT 180
181 AACATAGCTGTAGAACTACTAGGTGTATATAGGCGTACCGTTGCTGGACAAATAGCTAGT 240
181 AACATAGCTGTAGAACTACTAGGTGTATATAGGCGTACCGTTGCTGGACAAATAGCTAGT 240
241 TTTTATATAGTTTCTTTGTTGTTGTAATATATGCGCGCGAGAGATCAGTGGGAAATTTTC 300
241 TTTTATATAGTTTCTTTGTTGTTGTAATATATGCGCGCGAGAGATCAGTGGGAAATTTTC 300
301 CTAGAACATGTGCAACAACTTATAAATCAACAAATAACAGAAATGCTAGGAATACGGCA 360
301 CTAGAACATGTGCAACAACTTATAAAGACAAACAGTAAACAGAAATACTAGGGATACGGCT 360
361 CTGCTCGATTTACAAGGTTTAGGAGATTCCTTTAGAGCCCTATCAACAGTCACTTGAAGAT 420
361 CTGCTCGATTTACAAGGTTTAGGAGATTCCTTTAGAGCCCTATCAACAGTCACTTGAAGAT 420
421 TGGCTAGAAACCGTGATGATGCAAGAAACGAGAGTGTCTTTATACCCCAATATATAGCC 480
421 TGGCTAGAAACCGTGATGATGCAAGAAACGAGAGTGTCTTTATACCCCAATATATAGCC 480
481 TTAGAACTTGATTTCTTTAATGCGATGCGCGTTTTCGCAATTAAGAAACCAAGAGTTCCA 540
481 TTAGAACTTGATTTCTTTAATGCGATGCGCGTTTTCGCAATTAAGAAACCAAGAGTTCCA 540
541 TTATTATGTTATATGCTCAAGCTGCAAAATTTACACCTATTTATTTAGAGATGCTCT 600
541 TTATTATGTTATATGCTCAAGCTGCAAAATTTACACCTATTTATTTAGAGATGCTCT 600
601 CTTTTGTAGTGAATTTGGGCTTTACATCGCAGAAATTCAGGTTTATATGAGCGCAA 660
601 CTTTTGTAGTGAATTTGGGCTTTACATCCAGAAATTCAGGTTTATATGAGCGCAA 660
661 GTGGAAACAAACGAGAGATTTTCGACTATGCGTAGAATGGTATATAACAGTCTAAAT 720
661 GTGGAAACAAACGAGAGATTTTCGACTATGCGTAGAATGGTATATAACAGTCTAAAT 720
721 AGCTTGAGAGGGAACAAATGCGCAAGTTGGTGGTTATTAATCAATTCCTGAGAGACTA 780
721 AATTTGAGAGGGAACAAATGCTGAAAGTTGGTGGATTAATCAATTCCTGAGAGACTTA 780
781 ACCTTAGGGGTATTAGATCTAGTGGCACTATTCGCAAGCTATGACACTGCTTTATCCA 840
781 ACCTTAGGGGTATTAGATCTAGTGGCACTATTCGCAAGCTATGACACTGCTTTATCCA 840
841 ATAAATACGAGTGTCTAGTTAAACAGGGAAGTTTATACAGCGCAATTTGAGCAACAGGG 900
841 ATGAATACCAAGTGTCTCAATTTAAACAGGAAATTTATACAGATCCCAATTTGGGGAACAAAT 900
901 GTAAAT-----ATGCGCAAGTATGAATTTGGTATTAATAATATGACCTTCGTTTCGGCT 954
901 GCACCTTCAGAAATTTGCAAGTACGAATTTGGTATTAATAATATGACCTTCGTTTCGGCT 960
955 ATAGAGACTCGGGTTATTCGAAAGCCCGCATCTACTTGATTTTCTAGAAACAACTTACAAT 1014
961 ATAGAGCTCGCGTTATTAGGCTCCGCACTACTTGATTTTCTAGAAACAGCTTACAAT 1020
1015 TTTAGCACTTCAACGATGAGTGTCTAGGCACTATGAGGATATGACTTATGCGGGGGGACACA 1074
1021 TTAGCGGTATTAAGTTCGATGGAATTAATCAATATGAAATTAATGAGTGGTGGGACATAGA 1080
1075 ATTCAATCTCGGCCAATAGGAGCGGATTAATTAACCTCAACGATGGTCTACCAATATCT 1134
1081 CTGGAATCGCGAAACAAATAAGGGGTCAATTAAGTACCTTGGACACCGGAAATAACAAT 1140
1135 TCTATTAACTCTGTAGATTTATCACTTCTCTCGAGACGTTATTTGGACTGAATCATAT 1194
1141 TCTATTAACTCTGTAACTTACAGTTTCACTCTGAGACGTTTATGAAACAGAAATCATTT 1200
1195 GCAGGAGTGTCTTATGGGAAATTTACCTTGAACCTTATTCATGAGTGGTCCCTACTGTGTAGA 1254

1201 GCAGGAGTAATAATAT-----ACTTCTAACTACTCTCTGTGAATGGAGTACCTTGGGCTAGA 1254
1255 TTTAAATTTTAGGAACCTCTCAGAATATCTTTTGAAGAGGTACTGCTAACTATAGTCAACCC 1314
1255 TTTAAATTTGGAAGAAATCCCTGGAATTTCTTT---AGAGTTAGCCTTCTCTATATATAGG 1311
1315 TATGAGTCACTCTGGGCTTCAATTTAAAGATTCAGAACTGAAATTAACACAGAAACAAACA 1374
1312 TATACTGGAGTGGGACACAACTATTGTTTCAAGAACTGAAATTTACCACCAGAAACAAACA 1371
1375 GAACGACCAAAATTTATGAATCATATAGTATAGTATCTCACAATAGGGCTCATTTTCAAA 1434
1372 GAACGACCAAAATTTATGAATCTTACAGTCAATAGATTTATCTAATATAAGACTAAATAT 1431
1435 TCTAGGCTGATCTGACCAAGTATATTCTTGACGCCACCGTAGTCAGATCGTACAAATACC 1494
1432 AACACTTTGAGAGCAACAGTATATTCTTGAGCGCAACGTTAGTGCAGATCGTACAAATACC 1491
1495 ATTAGTTTCAGATAGCATAACACAAATACCAATTCGTAAATCATTTCAACCTTAAATTCAGGT 1554
1492 ATTAGTTTCAGATAGCATAACACAAATACCAATTCGTAAATCATTTCAACCTTAAATTCAGGT 1551
1555 ACCTCTGTAGTCACTGGCCCGAGGATTTACAGGAGGGGATATAATCCGAACTAAGTTAAT 1614
1552 ACCTCTGTAGTCACTGGCCCGAGGATTTACAGGAGGGGATATAATCCGAACTAAGTTAAT 1611
1615 GGTAGTGTACTAAGTATGGGTCTTAAATTTTAAATACATCATTTACAGCGGTATCGCGTG 1674
1612 GGTAGTGTACTAAGTATGGGTCTTAAATTTTAAATACATCATTTACAGCGGTATCGCGTG 1671
1675 AGATTTCTGTTATGCTCTTCTCAACAAATGGTCTCTGAGGGTAACTGTGCGAGGGAGTACT 1734
1672 AGATTTCTGTTATGCTCTTCTCAACAAATGGTCTCTGAGGGTAACTGTGCGAGGGAGTACT 1731
1735 ACTTTTGTCAAGGATTCCTTAGTACTATGAGTGCAGGAGTCTTTGACATCTCAATCA 1794
1732 ACTTTTGTCAAGGATTCCTTAGTACTATGAGTGCAGGAGTCTTTGACATCTCAATCA 1791
1795 TTTAGATTTTGCAAAATTTCTCTGTAGGTATTAGTGCATCTGCGAGTCAAACTGCTGGAATA 1854
1792 TTTAGATTTTGCAAAATTTCTCTGTAGGTATTAGTGCATCTGCGAGTCAAACTGCTGGAATA 1851
1855 AGTATAAGTAAATTAATGCAGGTAGACAAAGCTTTCACTTTGATATAAATTTGAATTCATTTCCA 1914
1852 AGTATAAGTAAATTAATGCAGGTAGACAAAGCTTTCACTTTGATATAAATTTGAATTCATTTCCA 1911
1915 ATTTACTGCAACCTTTGCAAGCAGGATACGATTTTGAAGGGCGCAAGGGCGGTGAATGCT 1974
1912 ATTTACTGCAACCTTTGCAAGCAGGATACGATTTTGAAGGGCGCAAGGGCGGTGAATGCT 1971
1975 CTGTTTACTAATAACGAATCCAAAGATTTGAAACAGATGTCAGAGATTTATCATATTGAT 2034
1972 CTGTTTACTAATAACGAATCCAAAGATTTGAAACAGGTTGCAAGATTTATCATATTGAT 2031
2035 CAAAGTATCCAAATTTAGTGGCGTGTATTATCGGATGAATTTCTGTTAGATGAAAGAGAGAA 2094
2032 GAAAGTATCCAAATTTAGTGGCGTGTATTATCGGATGAATTTCTGTTGATGAAAGAGAGAA 2091
2095 TTTACTTGAGAACTGAAATATGCGAAACGACTCAGTGTATGAAGAACTTACTTCCAGAT 2154
2092 TTTACTTGAGAACTGAAATATGCGAAACGACTCAGTGTATGAAGAACTTACTTCCAGAT 2151
2155 CCAAACTTCACATCCATTAAGCAACCAAGCTTCAATCTACTAATAGCAATCGAAT 2214
2152 CCAAACTTCACATCCATTAAGCAACCAAGCTTCAATCTACTAATAGCAATCGAAT 2211
2215 TTCACATCTATCCATGAACCAATCTGAACATGGATGGTGGGAAAGTGAGAACTTACAAATC 2274
2212 TTCACATCTATCCATGAACCAATCTGAACATGGATGGTGGGAAAGTGAGAACTTACAAATC 2271
2275 CAGGAAGGAATGACGTTATTTAAAGGAATTCAGTCACTACCGGGACTTTTAAATGAG 2334
2272 CAGGAAGGAATGACGTTATTTAAAGGAATTCAGTCACTACCGGGTACTTTTAAATGAG 2331

```
QY 2335 TGTATCCGACGTATTTATATCAAAAAATAGGAGAGTCGGAATTTAAAGCTTATACCTCGC 2394
DB 2332 TGTATCCGACGTATTTATATCAAAAAATAGGAGAGTCGGAATTTAAAGCTTATACCTCGC 2391
QY 2395 TACCAATTAAGAGGGTATATTGAAGATAGTCAAGATTTAGAGATATATTGATTCGTTAT 2454
DB 2392 TACCAATTAAGTGGCTATATTGAAGATAGTCAAGATTTAGAGATATATTGATTCGTTAT 2451
QY 2455 AATCGAAACATGAACATTTGGATGTTTCCAGGTACCGAGTCCGATGCCCGCTTTTCAGTT 2514
DB 2452 AATCGAAACATGAACATTTGGATGTTTCCAGGTACCGAGTCCGATGCCCGCTTTTCAGTT 2511
QY 2515 GAAAGCCCAATCGAAGGTGCGGAGAACCGAATCGATGCCACACATTTTGAATGGAAT 2574
DB 2512 GAAAGCCCAATCGAAGGTGCGGAGAACCGAATCGATGCCACACATTTTGAATGGAAT 2571
QY 2575 CCTGATCTAGATGTTCTCGCAGAGATGGAGAAAAATGTGGCATCATCTCCCATCATTTTC 2634
DB 2572 CCTGATCTAGATGTTCTCGCAGAGATGGAGAAAAATGTGGCATCATCTCCCATCATTTTC 2631
QY 2635 TCTTTGGATATTGATATTGGATGCACAGACTTGCATGAGAACTTAGCGGTGGGTGTTA 2694
DB 2632 TCTTTGGATATTGATATTGGATGCATGAGACTTGCATGAGAACTTAGCGGTGGGTGTTA 2691
QY 2695 TTCAGATTAAAGCGAGAGGTGATGCAAGACTAGGGAATCTGGAAATTTATTTGAAGAG 2754
DB 2692 TTCAGATTAAAGCGAGAGGTGATGCAAGACTAGGGAATCTGGAAATTTATTTGAAGAG 2751
QY 2755 AAACCATTTATAGAGAGCACTGCTCTGTGTGAAGAGACAGAGAAAAATGGAGAGAC 2814
DB 2752 AAACCATTTATAGAGAGCACTGCTCTGTGTGAAGAGACAGAGAAAAATGGAGAGAC 2811
QY 2815 AAACGTGAAAACTTACAAATTTGAAAAAAGAGTATATACAGAGGCAAAAGAGCTGTG 2874
DB 2812 AAACGTGAAAACTTACAAATTTGAAAAAAGAGTATATACAGAGGCAAAAGAGCTGTG 2871
QY 2875 GATGCTTTATTTGATGATTTCAATATTAATAGATTACAAGCGGATACAAACATTTGGCATG 2934
DB 2872 GATGCTTTATTTGATGATTTCAATATTAATAGATTACAAGCGGATACAAACATTTGGCATG 2931
QY 2935 ATTCATCGCGCAGATAACTGTTTCGATTCGAAATTCGAGAGCTTATCTGTCAGAAATATCT 2994
DB 2932 ATTCATCGCGCAGATAACTGTTTCGATTCGAAATTCGAGAGCGTATCTTTTCAGAAATATCT 2991
QY 2995 GTTATCCCGGTGTAAATGCGGAAATTTTGAAGATTTAGAAGTCCGATTCACCTGCA 3054
DB 2992 GTTATCCCGGTGTAAATGCGGAAATTTTGAAGATTTAGAAGTCCGATTCACCTGCA 3051
QY 3055 ATCTCCCTATACGATCGAGAAATGTCGTTTAAAAATGGTGATTTAATAATGGATTAGCA 3114
DB 3052 ATCTCCCTATACGATCGAGAAATGTCGTTTAAAAATGGTGATTTAATAATGGATTAGCA 3111
QY 3115 TGCTGGAATGTAAAGGGCATGTAGATGTACACAGAGCCATCCCGTTCGTGCTGTTGTT 3174
DB 3112 TGCTGGAATGTAAAGGGCATGTAGATGTGTACACAGAGCCATCCCGTTCGTGCTGTTGTT 3171
QY 3175 ATCCAGAAATGGGAAGCAGAGTGTCAACAGCAGTTCGCGTCTGTCGCGGCGCGTGGCTAT 3234
DB 3172 ATCCAGAAATGGGAAGCAGAGTGTCAACAGCAGTTCGCGTCTGTCGCGGCGCGTGGCTAT 3231
QY 3235 ATCTCCGTGTCAAGCGGTACAAAGAGGGATATGGAGAGGGTGTGTGTAAACGATCCATGAA 3294
DB 3232 ATCTCCGTGTCAAGCGGTACAAAGAGGGATATGGAGAGGGTGTGTGTAAACGATCCATGAA 3291
QY 3295 ATCGAGAAACATACAGACGAATTAATAATTTAAAACTGTGAAGAGAGGATGTATCCCA 3354
DB 3292 ATCGAGAAACATACAGACGAATTAATAATTTAAAACTGTGAAGAGAGGATGTATCCCA 3351
QY 3355 ACGGATACAGGACGTTGATGATTTATCTGCAACCAAGGTACAGAGTATGTATATCC 3414
DB 3352 ACGGATACAGGACGTTGATGATTTATCTGCAACCAAGGTACAGAGTATGTATATCC 3411
```

```
QY 3415 CGTAATCTCGATATGAGGATGCATATGAAGTTGATACTACAGCATCTGTAAATACAAA 3474
DB 3412 CGTAATCTCGATATGAGGATGCATATGAAGTTGATACTACAGCATCTGTAAATACAAA 3471
QY 3475 CCGACTTTATCAAGAAAGAAACGTATACAGATGTACGAAGAGATAATCATTTGTGAATATGAC 3534
DB 3472 CCGACTTTATCAAGAAAGAAACGTATACAGATGTACGAAGAGATAATCATTTGTGAATATGAC 3531
QY 3535 AGAGGGTATGTGAATTTATCCACCTACCAAGCTGTTTATATGACAAAGAAATTAAGATAC 3594
DB 3532 AGAGGGTATGTGAATTTATCCACCTACCAAGCTGTTTATATGACAAAGAAATTAAGATAC 3591
QY 3595 TTCCCGAAACCCGATAGGTATGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTG 3654
DB 3592 TTCCCGAAACCCGATAGGTATGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTG 3651
QY 3655 GACAGCGTGAATTTACTCTCTTATGGAGAAATAG 3687
DB 3652 GACAGCGTGAATTTACTCTTATGGAGAAATAG 3684

RESULT 7
AX098692 LOCUS 3684 bp DNA linear PAT 02-APR-2001
DEFINITION Sequence 62 from Patent WO0119859.
ACCESSION AX098692
VERSION AX098692.1 GI:13537934
KEYWORDS
SOURCE Bacillus thuringiensis
ORGANISM Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.

REFERENCE
1
AUTHORS Baum, J. A., Chu, C. R., Donovan, W. P., Gilmer, A. J. and Rupar, M. J.
TITLE Lepidopteran-active bacillus thuringiensis delta-endotoxin
JOURNAL compositions and methods of use
JOURNAL Patent: WO 0119859-A 62 22-MAR-2001;
JOURNAL MONSANTO COMPANY (US)
FEATURES
source
1. 3684
/mol_type="unassigned DNA"
/db_xref="taxon:1428"

ORIGIN
Query Match 88.4%; Score 3260.2; DB 6; Length 3684;
Best Local Similarity 93.3%; Pred. No. 0;
Matches 3445; Conservative 0; Mismatches 233; Indels 15; Gaps 3;

QY 1 TTGACTTCAAATAGGAAAAATGAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 60
DB 1 TTGACTTCAAATAGGAAAAATGAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 60
QY 61 TCGAATCATTTCCACACAAATGGATCTATCACCAGATGCTCGTATTGAGGATCTTTTGTGT 120
DB 61 TCGAATCATTTCCGACAAATGAAATCTATCAACCGATGCTCGTATTGAGGATGCTTGTGT 120
QY 121 ATAGCGAGGGGAAATATATCAATCCACTTTGTAGCGCATCAACAGTCCAAACCGGGTATT 180
DB 121 ATAGCGAGGGGAAACAAATATCGATCCATTTGTTAGCGCATCAACAGTCCAAACCGGGTATT 180
QY 181 AACATAGCTGGTGAATATCTAGGTGTATTAGGGGTACCGTTTGTGACAAATAGCTAGT 240
DB 181 AACATAGCTGGTGAATATCTAGGTGTATTAGGGGTACCGTTTGTGACAAATAGCTAGT 240
QY 241 TTTTATAGTTTTCTTGTGTGAAATATATGCGCCGCGGAGATCATAGTGGGAAATTTTC 300
DB 241 TTTTATAGTTTTCTTGTGTGAAATATATGCGCCGCGGAGATCATCTTGGGAAATTTTC 300
QY 301 CTAGAACATGTCGAACAACTTATAAATCAACAAATACAGAAATGCTAGGAATACCGCA 360
DB 301 CTAGAACATGTCGAACAACTTATAAAGCAACAAAGTAAACAGAAATACTAGGATACGGCT 360
```


QY 361 CTTGCTCGATTA CAAGGTTTAGGAGATTCTCTTAGAGCCTATCAACAGTCACCTTGAAGAT 420
 Db 361 CTTGCTCGATTA CAAGGTTTAGGAGATTCTCTTAGAGCCTATCAACAGTCACCTTGAAGAT 420
 QY 421 TGCGTAGAAGCCGTGATGATGCAAGAACGAGAGTGTCTTTATACCCCAATATATAGCC 480
 Db 421 TGCGTAGAAGAACCGTGAATGATGCAAGAACGAGAGTGTCTTTATACCCCAATATATAGCC 480
 QY 481 TTAGAACTTCGATTTCTTAATGCGATGCCGCTTTTCGCAATTTAGAAACCAAGAGTTCCA 540
 Db 481 TTAGAACTTCGATTTCTTAATGCGATGCCGCTTTTCGCAATTTAGAAACCAAGAGTTCCA 540
 QY 541 TTATTAATGATATATGCTCAAGTGCAGATTTTACACCTATTTATTTAGAGATGCTCT 600
 Db 541 TTATTAATGATATATGCTCAAGTGCAGATTTTACACCTATTTATTTAGAGATGCTCT 600
 QY 601 CTTTGTGATGATGAAATTTGGGCTTACATCGAGGAATTTCAACGTTATTTAGCGCCAA 660
 Db 601 CTTTGTGATGATGAAATTTGGGCTTACATCGAGGAATTTCAACGTTATTTAGCGCCAA 660
 QY 661 GTGGAACAAACGAGAGATTTTCGACTATTGCGTAGAATGGTATATACAGGTCTAAAT 720
 Db 661 GTGGAACAAACGAGAGATTTTCGACTATTGCGTAGAATGGTATATACAGGTCTAAAT 720
 QY 721 AGCTTAGAGGGAACAAATGCGCAAGTTGGTGTGCTTATTAATCAATTCGCTAGAGATCTA 780
 Db 721 AATTTAGAGGGAACAAATGCTGAAAGTTGGTGTGCTTATTAATCAATTCGCTAGAGATCTA 780
 QY 781 ACCTTAGGGGTATTTAGATCTAGTGGCACTATTTCCAGCTATGACACTCGCACTTATCCA 840
 Db 781 ACCTTAGGGGTATTTAGATCTAGTGGCACTATTTCCAGCTATGACACTCGCACTTATCCA 840
 QY 841 ATAAATACGAGTGTCTCAGTTTAAAGGGAAGTTTATACAGACGCAATTTGGAGCAACAGGG 900
 Db 841 ATGNAATACGAGTGTCTCAGTTTAAAGGGAAGTTTATACAGATCCAATTTGGAGCAACAA 900
 QY 901 GTAAAT-----ATGGCAAGTATGAATGGTATTAATTAATGAATGACCTTCGTTTCGGT 954
 Db 901 GCACCTTCAGGATTTGCAAGTACGAATTTGGTATTAATTAATGAATGACCACTTCGTTTCGGC 950
 QY 955 ATAGAGCTCGGCTTATCGAAGCCGCACTTACTTGAATTTCTAGAACCACTTGAATTT 1014
 Db 961 ATAGAGCTCGGCTTATTTAGGCTTCGCACTTACTTGAATTTCCAGAACAGCTTACAAAT 1020
 QY 1015 TTTAGCACTTTCATCAGATCGAGTGTCTTAGGATATGACTTACTTGGCGGGGCAACA 1074
 Db 1021 TTAGCGATTTAAGTCGATGGAGTAACTCAATATATGAATTAATTTGGTGGACATAGA 1080
 QY 1075 ATTCAATCTCGGCAATPAGAGCGGATTAATTAACCTCAACGATGGGTCTACCAATFACT 1134
 Db 1081 CTTGAATCGGCAACAAATAGGGGTCTAATTAAGTACCTGGACACACGGAATAACCAATFACT 1140
 QY 1135 TCTATTAATCTGTAGATTTATCACTTCTTCGAGACGCTATATTTGGACTGAATCATAT 1194
 Db 1141 TCTATTAATCTGTAAATCAATTTACAGTTTCACTTCGAGACGCTTATATGAACAGAAATCATTT 1200
 QY 1195 GCAGGAGTGTCTTATGCGGAAATTTACCTTGAACCTATTTCACTGTGCTCCTACTGTTAGA 1254
 Db 1201 GCAGGGAATAATAT-----ACTTCTAATCTACTCTGTGAATGAGTACCTTGGCTAGA 1254
 QY 1255 TTTAAATTTTAGAACCTCTAGAAATPACTTTTGAAGAGGTACTGCTAACTATAGTCAACCC 1314
 Db 1255 TTTAAATTTAGAAATCCCTGAAATCTCTT---AGAGGTAGCTTCTCTATATATAGGG 1311
 QY 1315 TATGAGTCACTCGGCTTCAATTTAAAGATTCAGAACTGAATTTACCACCAAGAACAA 1374
 Db 1312 TATACCTGGAGTGGGACACAACTATTTGATTCAGAACTGAATTTACCACCAAGAACAA 1371
 QY 1375 GAACGACCAAAATATGAATCATATAGTCAATAGGTTATCTCACTAGGGCTCATTTTCAAA 1434
 Db 1372 GAACGACCAAAATATGAATCTTACAGTCATAGATTATCTAATATAGCAATATATCAGGA 1431
 QY 1435 TCTAGGTCATGTACCAAGTATATTTCTTGGACGACCGTAGTCAGATCGTACAAATACC 1494

Db 1432 AACACTTTGAGACCAACAGTATATCTTGGACCAACCGTAGTCAGATCGTACAAATACC 1491
 QY 1495 ATTAGTTTCAGATAGCATAAACAAATACCAATTCGTTAAATCATTTCAACCTTTAATTCAGGT 1554
 Db 1492 ATTAGTTTCAGATAGCATAAACAAATACCAATTCGTTAAATCATTTCAACCTTTAATTCAGGT 1551
 QY 1555 ACCTCTGTAGTCAGTGGCCAGGATTTTACAGGAGGGGATATAATCCGAACTAACGTTAAT 1614
 Db 1552 ACCTCTGTAGTCAGTGGCCAGGATTTTACAGGAGGGGATATAATCCGAACTAACGTTAAT 1611
 QY 1615 GGTAGTGTACTAAGATGAGTGGTCTTAATTTTAAATAATACATCATTTACAGCGGTATCGCGTG 1674
 Db 1612 GGTAGTGTACTAAGATGAGTGGTCTTAATTTTAAATAATACATCATTTACAGCGGTATCGCGTG 1671
 QY 1675 AGATTTGCTGTATCTCTCTCAACAAATGCTCTGAGGGTAACTGTCTCGAGGAGTACT 1734
 Db 1672 AGATTTGCTGTATCTCTCTCAACAAATGCTCTGAGGGTAACTGTCTCGAGGAGTACT 1731
 QY 1735 ACTTTTGTATCAAGGATTTCCCTAGTACTATGAGTGCAAATGAGTCTTTTGACATCTCAATCA 1794
 Db 1732 ACTTTTGTATCAAGGATTTCCCTAGTACTATGAGTGCAAATGAGTCTTTTGACATCTCAATCA 1791
 QY 1795 TTTAGATTTTCAGAAATTTCCCTGTAGTATTTAGTGCATCTGGCAGTCAAACCTGCTGGAATA 1854
 Db 1792 TTTAGATTTTCAGAAATTTCCCTGTAGTATTTAGTGCATCTGGCAGTCAAACCTGCTGGAATA 1851
 QY 1855 AGTATAAGTAAATTAATGCAAGTACAAACGTTTTCATTTGATAAATTTGAATTTCAITCCA 1914
 Db 1852 AGTATAAGTAAATTAATGCAAGTACAAACGTTTTCATTTGATAAATTTGAATTTCAITCCA 1911
 QY 1915 ATTTACTGCAACCTTCGAAGCAGAAATACGAATTTTGAAGGGCGCAAGGGCGGTGAATGCT 1974
 Db 1912 ATTTACTGCAACCTTCGAAGCAGAAATATGATTTTGAAGAGCGCAAGGGCGGTGAATGCT 1971
 QY 1975 CTGTTTACTTAATACGAATTCGAAGAGATTTGAACACAGATGTGCAGATTTATCATATTGAT 2034
 Db 1972 CTGTTTACTTAATACGAATTCGAAGAGGTTTGAACACAGGTTGTGCAGATTTATCATATTGAT 2031
 QY 2035 CAAGTATCCAAATTTAGTGGCGTCTTTATCGATGAATTTCTGCTTAGATGAAGAGAGAA 2094
 Db 2032 GAAGTATCCAAATTTAGTGGCGTCTTTATCGATGAATTTCTGCTTAGATGAAGAGAGAA 2091
 QY 2095 TTAATTTGAGAAAGTGAATAATGCGAAACGACTCAGTGTGATGAAGAACTTTACTCCAAGAT 2154
 Db 2092 TTAATTTGAGAAAGTGAATAATGCGAAACGACTCAGTGTGATGAAGAACTTTACTCCAAGAT 2151
 QY 2155 CCAAACTTCAATCCATCAATTAAGCAACCAAGCTTCAATTTCTAATTAATGAGCAATCGAAT 2214
 Db 2152 CCAAACTTCAATCCATCAATTAAGCAACCAAGCTTCAATTTCTAATTAATGAGCAATCGAAT 2211
 QY 2215 TTTCAATCTATCCATCAATCTGAACCAATCTGGAAGTGGTGGGAAAGTGAAGCAATTTACAATC 2274
 Db 2212 TTTCAATCTATCCATCAATCTGAACCAATCTGGAAGTGGTGGGAAAGTGAAGCAATTTACAATC 2271
 QY 2275 CAGGAAGGAATCAACGTTATTTAAAGAAATTCAGTCACTACCGGGGACTTTTAAATGAG 2334
 Db 2272 CAGGAAGGAATCAACGTTATTTAAAGAAATTCAGTCACTACCGGGGACTTTTAAATGAG 2331
 QY 2335 TGTATTCGAGCTATTTATTAATCAAAAAATAGGAGAGTCCGAAATTAAGCTTATACCTGC 2394
 Db 2332 TGTATTCGAGCTATTTATTAATCAAAAAATAGGAGAGTCCGAAATTAAGCTTATACCTGC 2391
 QY 2395 TACCAATTTAAGGGGTATATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT 2454
 Db 2392 TACCAATTTAAGGGGTATATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT 2451
 QY 2455 AATGCGAAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTCAAGT 2514
 Db 2452 AATGCGAAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTCAAGT 2511
 QY 2515 GAAGGCCAATCCGAAGGTGCGGAGAACCGAATTCGATGCGCACCAATTTGATGGAAT 2574

2512 GAAAGCCCAATCGAAGGTCGGAGAACCGAATCGATCGGCACACACATTTTGAATGGAAT 2571
 2575 CCGTATCTAGATGTTCTCTCGAGAGATGGAAGAAATGTGCGCATCATTCCTCCATCATTT 2634
 2572 CCGTATCTAGATGTTCTCTCGAGAGATGGAAGAAATGTGCGCATCATTCCTCCATCATTT 2631
 2635 TCCTTGGATATTTGATTTGATGTCACAGACTTGTGATGAGATCTAGCGGTGTGGTGGA 2694
 2632 TCCTTGGATATTTGATTTGATGTCACAGACTTGTGATGAGATCTAGCGGTGTGGTGGA 2691
 2695 TTCAAGATTAAAGACGCGAGGAGGTCATGCAAGACTTAGGGAATCTGGAATTTATTGAAGAG 2754
 2692 TTCAAGATTAAAGACGCGAGGAGGTCATGCAAGACTTAGGGAATCTGGAATTTATTGAAGAG 2751
 2755 AAACCAATTATTAGAGAGACCTGTCTGTGTGAAGAGACAGAGAAAAAATGGAGAGAC 2814
 2752 AAACCAATTATTAGAGAGACCTGTCTGTGTGAAGAGACAGAGAAAAAATGGAGAGAC 2811
 2815 AAACGTGAAAACCTACAAATTGGAACAAAACGAGTATATACAGAGGCAAAAGAGCTGTG 2874
 2812 AAACGTGAAAACCTACAAATTGGAACAAAACGAGTATATACAGAGGCAAAAGAGCTGTG 2871
 2875 GATGCTTTATTGTAGATTTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATG 2934
 2872 GATGCTTTATTGTAGATTTCTCAATATAGATTACAAGCGGATACAAACATTTGGCATG 2931
 2935 ATTCAATCGGAGATAAATCTGTTTCATGAAATTCAGAGGCTTATCTGTGAGAAATATCT 2994
 2932 ATTCAATCGGAGATAAATCTGTTTCATGAAATTCAGAGGCTTATCTGTGAGAAATATCT 2991
 2995 GTTATCCCGGTGTAAATGCGGAATTTTGAAGAATTAGAGGTCGATATCACTGCA 3054
 2992 GTTATCCCGGTGTAAATGCGGAATTTTGAAGAATTAGAGGTCGATATCACTGCA 3051
 3055 ATCTCCCTATACGATGCGAGAAATGTCGTTAAATGCTGATTTAATGATGATGCA 3114
 3052 ATCTCCCTATACGATGCGAGAAATGTCGTTAAATGCTGATTTAATGATGATGCA 3111
 3115 TGCTGGAATGTAAAGGCGATGTAGATGTACAAAGAGCCATCACCGTCTGCTGCTGTT 3174
 3112 TGCTGGAATGTAAAGGCGATGTAGATGTACAAAGAGCCATCACCGTCTGCTGCTGTT 3171
 3175 ATCCAGATGCGGAGAGAGATGTACAGAGTTCGCTGCTGCTGCGGGCGTGGCTAT 3234
 3172 ATCCAGATGCGGAGAGAGATGTACAGAGTTCGCTGCTGCTGCGGGCGTGGCTAT 3231
 3235 ATCTCCGCTGTACAGAGTCAAGAGGAGATGTGAGAGGCTTCTGTAAACGATCATGAA 3294
 3232 ATCTCCGCTGTACAGAGTCAAGAGGAGATGTGAGAGGCTTCTGTAAACGATCATGAA 3291
 3295 ATCCAGAACCAATACAGACGAATTTAAATTTAAACCTGTGAGAGAGGAGTGTATCCA 3354
 3292 ATCCAGAACCAATACAGACGAATTTAAATTTAAACCTGTGAGAGAGGAGTGTATCCA 3351
 3355 ACAGATACAGAAAGCTGATGATATATCTGACACCAAGGATACAGAGTATGTAATCC 3414
 3352 ACAGATACAGAAAGCTGATGATATATCTGACACCAAGGATACAGAGTATGTAATCC 3411
 3415 CGTAACTCTGATATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3474
 3412 CGTAACTCTGATATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3471
 3475 CCGACTTATGAAGAGAGAAAGCTATACAGATGTACAGAGAGATATCAATTTGTAATATGAC 3534
 3472 CCGACTTATGAAGAGAGAAAGCTATACAGATGTACAGAGAGATATCAATTTGTAATATGAC 3531
 3535 AGAGGATATGTAATTTATCCACCTACCTAGCTGTTTATATGACAAAGAAATTTAGATATC 3594
 3532 AGAGGATATGTAATTTATCCACCTAGCTGTTTATATGACAAAGAAATTTAGATATC 3591
 3595 TTCCAGAAACCGATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3654
 3592 TTCCAGAAACCGATAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3651

Qy 3655 GACAGCGTGAATTAATCTCTTATGGAGGAATAG 3687
 Db 3652 GACAGCGTGAATTAATCTCTTATGGAGGAATAG 3684

RESULT 8

BTCRYA4
 LOCUS B. thuringiensis cryA4 gene for delta-endotoxin.
 DEFINITION
 ACCESSION X06711.1 GI:40264
 VERSION
 KEYWORDS cryA4 gene; crystal protein; delta-endotoxin; endotoxin.
 SOURCE Bacillus thuringiensis
 ORGANISM Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.

REFERENCE

1 (bases 1 to 3687)
 Brizzard, B.L. and Whiteley, H.R.
 Nucleotide sequence of an additional crystal protein gene cloned from *Bacillus thuringiensis* subsp. *thuringiensis*
 Nucleic Acids Res. 16 (6), 2723-2724 (1988)
 88203216
 3362680

REFERENCE

2 (bases 1 to 3687)
 Whiteley, H.R.
 Direct Submission
 Submitted (26-JAN-1988) Whiteley H. R., Department of W, Department of Microbiology, University of Washington, Seattle, WA 98195
 JOURNAL Location/Qualifiers

FEATURES

source
 1..3687
 /organism="Bacillus thuringiensis"
 /mol_type="genomic DNA"
 /strain="HD-2"
 /db_xref="taxon:1428"
 1..3687
 /notes="unnamed protein product; delta-endotoxin (AA 1-1228)"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA29898.1"
 /db_xref="GI:580949"
 /db_xref="GOA:P05517"
 /db_xref="UniProt/Swiss-Prot:P05517"
 /translation="MTSNRKNEIIINAVNSHAQMDLPDARIEDSLCIAEAGNNIDP
 FVSAITVQGINIAGRIILVGLVGFAGQSLAFYFLVGLWPRGRDQWEIFLEHVEQL
 INQITENARNTALRLQGLDSFRAYQSLDLENDNRDARTSVLYTOVIALDELDF
 LNAMPFLAPNOEVLPLMVAQANLHLLRLDASLFGSEFGLTSQIORYVERQVER
 TRYSDICVENYNTGLNSLGTNAASWRVNRQRRDLTLGLDLVALFPYSDTYPI
 NTSQUTREVTYDAIGATGVNMAWMNNNAPFSAIAEAAKSPHLLDFLEQITIF
 SASRSNTRHMTYWRGHTIQSRPIGGGLNTSTHGATNTSINPVTLFASRDVYRTES
 YAGVLLGIYLEPIHGVTFRFTNPQNI SDRGATANYSPQYSGIQLKDSYTELPP
 ETTERPNYESYSHRLSHIGIILQSRVNVVYVSWTHRSADRTNTIGPNRIQIIPMKVAS
 ELPGQTVVVGPGTGDILRRNTGGPGPIRVTVNGPLTQRVIGRFRYASTVDPDFE
 VSGGTTVNNFRFLRWNSGDELKYGNFVRRATPTPTFTQIDIIIRTSIQGLSGNE
 VYIDKLEIIPVATFAEYDLERQAVNALFTNPNRUKTDIVDTHIDQVSNLVC
 QSFHGWGSENIITQEGNDVFENYVTLPGTFNECYTLYIKQIGESLKAFTYQLR
 GYEDSODLEIYILRYNAKHETLDVPGTESLPLSPVESPIGRCGEPNRCAPHPFENPD
 LDCSGDEKCAHSHHSLDIDVCTDLHENLGVWVFKIKTQEGHARLGNLFIEE
 KPLGALSRLKRAEAKWRDKRKLQLETXYRTEAKEAVDALFPVDSQYDLQADTNI
 KTHADKLVHRIRREALYSELPIPGVNAIEFELEGHIIITALSILDARNVVRKNGDFN
 NGITCNWGHVDVQSHHRSDLVPEWEAEVQAVRVCPCGGLYLRVATYKEGYGEG
 CVTHIETNNDELKFTKREBEVYPTDTCNDYTAHQGTAGCADCACNRNAGYBDA
 YEVDTTASVNYKTYBEETVTDVRRDNHCEYDRGVNYPVPVAGYVTKLEYFPETDT
 WIEIGETGKFIQVDSVELLMEE"

CDS

1..3687
 /notes="unnamed protein product; delta-endotoxin (AA 1-1228)"
 /codon_start=1
 /transl_table=11
 /protein_id="CAA29898.1"
 /db_xref="GI:580949"
 /db_xref="GOA:P05517"
 /db_xref="UniProt/Swiss-Prot:P05517"
 /translation="MTSNRKNEIIINAVNSHAQMDLPDARIEDSLCIAEAGNNIDP
 FVSAITVQGINIAGRIILVGLVGFAGQSLAFYFLVGLWPRGRDQWEIFLEHVEQL
 INQITENARNTALRLQGLDSFRAYQSLDLENDNRDARTSVLYTOVIALDELDF
 LNAMPFLAPNOEVLPLMVAQANLHLLRLDASLFGSEFGLTSQIORYVERQVER
 TRYSDICVENYNTGLNSLGTNAASWRVNRQRRDLTLGLDLVALFPYSDTYPI
 NTSQUTREVTYDAIGATGVNMAWMNNNAPFSAIAEAAKSPHLLDFLEQITIF
 SASRSNTRHMTYWRGHTIQSRPIGGGLNTSTHGATNTSINPVTLFASRDVYRTES
 YAGVLLGIYLEPIHGVTFRFTNPQNI SDRGATANYSPQYSGIQLKDSYTELPP
 ETTERPNYESYSHRLSHIGIILQSRVNVVYVSWTHRSADRTNTIGPNRIQIIPMKVAS
 ELPGQTVVVGPGTGDILRRNTGGPGPIRVTVNGPLTQRVIGRFRYASTVDPDFE
 VSGGTTVNNFRFLRWNSGDELKYGNFVRRATPTPTFTQIDIIIRTSIQGLSGNE
 VYIDKLEIIPVATFAEYDLERQAVNALFTNPNRUKTDIVDTHIDQVSNLVC
 QSFHGWGSENIITQEGNDVFENYVTLPGTFNECYTLYIKQIGESLKAFTYQLR
 GYEDSODLEIYILRYNAKHETLDVPGTESLPLSPVESPIGRCGEPNRCAPHPFENPD
 LDCSGDEKCAHSHHSLDIDVCTDLHENLGVWVFKIKTQEGHARLGNLFIEE
 KPLGALSRLKRAEAKWRDKRKLQLETXYRTEAKEAVDALFPVDSQYDLQADTNI
 KTHADKLVHRIRREALYSELPIPGVNAIEFELEGHIIITALSILDARNVVRKNGDFN
 NGITCNWGHVDVQSHHRSDLVPEWEAEVQAVRVCPCGGLYLRVATYKEGYGEG
 CVTHIETNNDELKFTKREBEVYPTDTCNDYTAHQGTAGCADCACNRNAGYBDA
 YEVDTTASVNYKTYBEETVTDVRRDNHCEYDRGVNYPVPVAGYVTKLEYFPETDT
 WIEIGETGKFIQVDSVELLMEE"

ORIGIN

Query Match 85.1%; Score 3136.8; DB 1; Length 3687;
 Best Local Similarity 91.2%; Pred. No. 0;
 Matches 3375; Conservative 0; Mismatches 297; Indels 30; Gaps 3;
 Qy 1 TTGACTCAATATAGGAAAAATGAGATGAATTAATAATGCTTTATCGATTCGACTGTA 60

|||||
1 TTGACTTCAAAATAGGAAAAATGAAATGAAATATAAAAT-----GCTGTA 45
61 TCGAATCATTTCCACAAAAATGGATCTATCACCAGATGCTCGTATTTGAGGATTTCTTTGTTG 120
46 TCGAATCATTTCCGACAAAATGGATCTATTAACAGATGCTCGTATTTGAGGATAGCTTTGTTG 105
121 ATAGCCGAGGGGAATAATATCAATCCACTTGTAGCGCATCAACAGTCCCAACGGGTATT 180
106 ATAGCCGAGGGGAACAATATCGATCCATTTGTTAGCGCATCAACAGTCCCAACGGGTATT 165
181 AACATAGCTGGTAGAATACTAGTGTAATTAAGCGGTACCGTTTGTCTGCAAAAATAGCTAGT 240
166 AACATAGCTGGTAGAATACTAGGCGTATTTGGGCGTACCGTTTGTCTGCAACAATAGCTAGT 225
241 TTTTATAGTTTTCTTGTGTGTGAATTAATGCGCGCGGCGGAGATCAGTGGGAAATTTTC 300
226 TTTTATAGTTTTCTTGTGTGTGAATTAATGCGCGCGGCGGAGATCAGTGGGAAATTTTC 285
301 CTAGAAATGTCGAAACAACTTATAAAATCAACAAATAACAGAAATGCTAGGAATACGGCA 360
286 CTAGAAATGTCGAAACAACTTATAAAATCAACAAATAACAGAAATGCTAGGAATACGGCT 345
361 CTGCTCGATTAACAGGTTTAGAGATTCCTTTAGAGCTTATCAACAGTCACTTTGAAGAT 420
346 CTGCTCGATTAACAGGTTTAGAGATTCCTTTAGAGCTTATCAACAGTCACTTTGAAGAT 405
421 TGGCTAGAAACCGTGATGATGAAGAACGAGAGTGTCTTTATACCCCAATATATAGCC 480
406 TGGCTAGAAACCGTGATGATGAAGAACGAGAGTGTCTTTATACCCCAATATATAGCT 465
481 TTAGAACTTGATTTTCTTAATGCGATCGCGCTTTTCGCAATTAGAAACCAAGATTCCA 540
466 TTAGAACTTGATTTTCTTAATGCGATCGCGCTTTTCGCAATTAGAAACCAAGATTCCA 525
541 TTATTAATGTTATGCTCAAGCTGCAAAATTTACACCTATTTATTTAGAGATGCTCT 600
526 TTATTAATGTTATGCTCAAGCTGCAAAATTTACACCTATTTATTTAGAGATGCTCT 585
601 CTTTGTGTTAGTGAATTTGGGCTTACATCGCAGAAATTCACGTTTATTTAGCGCCAA 660
586 CTTTGTGTTAGTGAATTTGGGCTTACATCGCAGAAATTCACGTTTATTTAGCGCCAA 645
661 GTGGAAACAAACGAGAGATTTCCGACTATTTGGTGAATTTGTTATATAACAGGTCTAAAT 720
646 GTGGAAACAAACGAGAGATTTCCGACTATTTGGTGAATTTGTTATATAACAGGTCTAAAT 705
721 AGCTTGAGAGGACAAATGCGCGAGTTGGTGGTTTATATCAATCCATTCGCTAGATCTTA 780
706 AGCTTGAGAGGACAAATGCGCGAGTTGGTGGTTTATATCAATTCGCTAGAGATCTTA 765
781 ACGTTAGGGTATTAGATCTAGTGGCACTATTCGAGCTATGACACTCGCCTTTATCCA 840
766 ACGTTAGGGTATTAGATCTAGTGGCACTATTCGAGCTATGACACTCGCCTTTATCCA 825
841 ATAAATACGAGTGTCTAGTTTAAACAGGGAAGTTTATACAGACGCAATTTGAGCAACAGG 900
826 ATAAATACGAGTGTCTAGTTTAAACAGGGAAGTTTATACAGACGCAATTTGAGCAACAGG 885
901 GTAAATATGCAAGTATGAATTTGGTATTAATAATATGCACTTCGTTTCCGCTATAGAG 960
886 GTAAATATGCAAGTATGAATTTGGTATTAATAATATGCACTTCGTTTCCGCTATAGAG 945
961 ACTGCGGTATTCGAGCCGCACTACTTTGATTTTCTAGAACAACTTACAAATTTTATG 1020
946 GCTGCGGTATTCGAGCCGCACTACTTTGATTTTCTAGAACAACTTACAAATTTTATG 1005
1021 ACTTCATCAGATGGAGTGTCTAGGCAATATGACTTACTGCGGGGGGACAAATTTCAA 1080
1006 GCTTCATCAGATGGAGTGTCTAGGCAATATGACTTACTGCGGGGGGACAAATTTCAA 1065
1081 TCTCGGCAATAGAGCGGATTAATACCTCAACGATGGGTCTACCAATATCTTATT 1140
|||||

Db
1066 TCTCGGCCAATAGGAGCGGATTTAAATACTCAACGATGGGCTACCAATCTTCTTATT 1125
QY
1141 AATCCTGTAAGATTATCATTTCTCTCGAGAGCTATATTGGACTGAATCATATGACAGGA 1200
Db
1126 AATCCTGTAACATTACGTTCCGATCTCGAGACGTTTATAGGACTGAATCATATGACAGGA 1185
QY
1201 GTGCTTCTATGGGGAATTTACTTTGAACCTTATTCATGGTGTCCCTACTGTTAGATTAAAT 1260
Db
1186 GTGCTTCTATGGGGAATTTACTTTGAACCTTATTCATGGTGTCCCTACTGTTAGTTAAAT 1245
QY
1261 TTTAGGAACCTTCAGAAATCTTTTGAAGAGGTACTGCTAACTATATAGTCAACCTATGAG 1320
Db
1246 TTTACGAACCTTCAGAAATATTTCTGATAGAGGTACCGCTAACTATATAGTCAACCTATGAG 1305
QY
1321 TCACCTGGGCTTCAATTTAAAGATTCAGAAATTCAGAAATTCACACAGAAACAAACAGAACGA 1380
Db
1306 TCACCTGGGCTTCAATTTAAAGATTCAGAAATTCAGAAATTCACACAGAAACAAACAGAACGA 1365
QY
1381 CCAAAATATGAATCATATAGTCTATAGTGTATCTCATATAGGCTCAATTTTCAAACTCTAGG 1440
Db
1366 CCAAAATATGAATCTTTACAGTCAAGGTTATCTCATATAGGCTCAATTTTCAAACTCTAGG 1425
QY
1441 GTGCATGTACCAAGTATATTTCTGGACGCAAGGTAGTGCAGATCGTACAAATACCATAGT 1500
Db
1426 GTGAATGTACCGGTATATTTCTGGACGCAAGGTAGTGCAGATCGTACAAATACCATAGT 1485
QY
1501 TCAGATAGCATACAAATACCAATTCGTTAAATCATTCAACTTAAATTCAGGTACTCT 1560
Db
1486 CCAAAATAGAAATCACCCAAATCCCAATGGTAAAGCATTCGAATTTCTCAAGGTACCACT 1545
QY
1561 GTAGTCAGTGGCCAGGATTTACAGAGGGATATTAATCCGAACTAAACGTTAAATGGTAGT 1620
Db
1546 GTTGTGTAGAGGACAGGATTTTACTGTTGGGATATTTCTTGAAGAACGAATATCTGGTGA 1605
QY
1621 GTACTAAGTATGGTCTTAAATTTTAAATATACATCATTCAGAGCGGTATCCGCTGAGATT 1680
Db
1606 TTTGGACCGGATAAGAGTAACTGTATAACGACCACTAAACAAAGATATCGTATAGGATTC 1665
QY
1681 CGTTATGCTGCTTCTCAACAAATGGTCTCGAGGGTAACTGTTCGAGGGAGTACTACTTTT 1740
Db
1666 CGCTATGCTTCACTGTAGATTTTGATTTCTTTGTATCAGTGGAGGTACTACTGTAAAT 1725
QY
1741 GATCAAGGATTCCTTAGTACTATGATGCAAAATGAGTCTTTTGACATCTCAATCATTTAGA 1800
Db
1726 AATTTTAGATTTCTAGCTCAATGAACAGTGGAGACGAATAAAATACGGAATTTTGTG 1785
QY
1801 TTTGCAAGATTTCTGTAGTATTTAGTGCATCTGGCAGTCAA--ACTGCTGCAATAAGT 1857
Db
1786 AGACGTGCTTTTACTACACCTTTTACTTTTACACAAATTCAGATATAATTCGAACGTCT 1845
QY
1858 ATAAGTAATAATGCAGGTAGACAAACGTTTCACTTTGATAAAATTCGAATTCATTTCAAAT 1917
Db
1846 ATTCAGAGGCTTAGTGGAAATGGGGAAGTCTATATAGATAAAATTCGAATTTATCCAGTT 1905
QY
1918 ACTGCAACCTTGAAGCAAGATACGATTTAGAAAGGGCGCAAGCGGTGGAATGCTCTG 1977
Db
1906 ACTGCAACCTTGAAGCAAGATACGATTTAGAAAGAGCGCAAGCGGTGGAATGCTCTG 1965
QY
1978 TTTTAAATACGAAATCCAGAGATTTGAAACAGATGTCAGATTTATCATTTGATCAA 2037
Db
1966 TTTTAAATACGAAATCCAGAGATTTGAAACAGATGTCAGATTTATCATTTGATCAA 2025
QY
2038 GTATCCAAATTTAGTGGCTGTTTATCGGATGAATTTCTGTTAGATGAAAGAGAGAAATTA 2097
Db
2026 GTATCCAAATTTAGTGGCTGTTTATCGGATGAATTTCTGTTGATGAAAGAGAGAAATTA 2085
QY
2098 CTTGAGAAATGGAATATGCGAAACGACTCAGTGATGAAAGAACTTACTCCAGATCCA 2157
Db
2086 CTTGAGAAATGGAATATGCGAAACGACTCAGTGATGAAAGAACTTACTCCAGATCCA 2145
QY
2158 AACTTCACATCCATCAATAGGCAACGACTTTCATATCTACTAATGAGCAATCGAATTTTC 2217
Db
2146 AACTTCACATCCATCAATAGGCAACGACTTTCATATCTACTAATGAGCAATCGAATTTTC 2205

```
Qy 2218 ACATCTATCCATGAACAACTCTGAACATCGATGGTGGGGAAGTGAGAACATTAACAATCCAG 2277
Db 2206 ACATCTATCCATGAACAACTCTGAACATCGATGGTGGGGAAGTGAGAACATTAACAATCCAG 2265
Qy 2278 GAAGGAATGACGTATTTAAAGAGAAATTAAGTCACTACCTACCGGGGACTTTTAAATGAGTGT 2337
Db 2266 GAAGGAATGACGTATTTAAAGAGAAATTAAGTCACTACCTACCGGGGACTTTTAAATGAGTGT 2325
Qy 2338 TATCCGACGTATTTATATCAAAAAATAGGAGAGTCGGAATTTAAAGCTTATACCTCGCTAC 2397
Db 2326 TATCCGACGTATTTATATCAAAAAATAGGAGAGTCGGAATTTAAAGCTTATACCTCGCTAC 2385
Qy 2398 CAATTAGAGGGTATATTAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTATAAT 2457
Db 2386 CAATTAGAGGGTATATTAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTATAAT 2445
Qy 2458 GCGAAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTTCAGTTGAA 2517
Db 2446 GCGAAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTTCAGTTGAA 2505
Qy 2518 AGCCCAATCGGAAGTTCGGAGAACCGAATCGATCGCACACATTTTGAATGGAATCCT 2577
Db 2506 AGCCCAATCGGAAGTTCGGAGAACCGAATCGATCGCACACATTTTGAATGGAATCCT 2565
Qy 2578 GATCTAGATTGTTCTCGCAGAGATGGAAGAAAAATGTCGCGATCATTTCCCATCAATTTCTCT 2637
Db 2566 GATCTAGATTGTTCTCGCAGAGATGGAAGAAAAATGTCGCGATCATTTCCCATCAATTTCTCT 2625
Qy 2638 TTGGATATTCATATTTGATGCACAGACTTCATGAGATCTAGCGGTGGTGGTATTC 2697
Db 2626 TTGGATATTCATATTTGATGCACAGACTTCATGAGATCTAGCGGTGGTGGTATTC 2685
Qy 2698 AAGATTAAAGACGACGAGGAGTCTAGCAAGACTAGGGAATCTGGAATTTATTTGAAGAGAAA 2757
Db 2686 AAGATTAAAGACGAGGAGGAGTCTAGCAAGACTAGGGAATCTGGAATTTATTTGAAGAGAAA 2745
Qy 2758 CCATTATTAGGAGAGCACTGTTCTGTTGGAAGAGACAGAGAAAAAATCGAGAGACAAA 2817
Db 2746 CCATTATTAGGAGAGCACTGTTCTGTTGGAAGAGGCGAGAGAAAAAATCGAGAGACAAA 2805
Qy 2818 CGTGAACAACTCAATTTGGAACAAACAGAGTATATACAGAGGCAAAAGAGCTGGAT 2877
Db 2806 CGTGAACAACTCAATTTGGAACAAACAGAGTATATACAGAGGCAAAAGAGCTGGAT 2865
Qy 2878 GCTTTATTTCTAGATTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATGATT 2937
Db 2866 GCTTTATTTCTAGATTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATGATT 2925
Qy 2938 CATCGCGCAGATAAACTTTGTTTCATCGAATTCGAGAGCTTATCTGTGAGAAATTTATCTGTT 2997
Db 2926 CATCGCGCAGATAAACTTTGTTTCATCGAATTCGAGAGCTTATCTTTTCAGAAATTTACCTGTT 2985
Qy 2998 ATCCGGGTGTAATTCGGGAATTTTGAAGATTTAGAAAGTCGATATACACTGCATC 3057
Db 2986 ATCCCGAGGTGTAATTCGGGAATTTTGAAGATTTAGAAAGTCACATATACACTGCATC 3045
Qy 3058 TCCTCTATACGATCGCAGAAATGTCGTTTAAATAATGATTTAATATGATTTAGATGATC 3117
Db 3046 TCCTCTATACGATCGCAGAAATGTCGTTTAAATAATGATTTAATATGATTTAATGATTTAATGAT 3105
Qy 3118 TGGAAATGTAAGGCGCATGTAGATGTACAACAGAGCCATCACCTGTTCTGTTGTTATC 3177
Db 3106 TGGAAATGTAAGGCGCATGTAGATGTACAACAGAGCCATCATCGTTCTGACCTGTTGTTATC 3165
Qy 3178 CAGAAATGGAAGCAGAGTGTCAAGAGAGTTCGCGTCTGTCGGGGCGTGGCTATATC 3237
Db 3166 CAGAAATGGAAGCAGAGTGTCAAGAGAGTTCGCGTCTGTCGGGGCGTGGCTATATC 3225
Qy 3238 CTCGGTGTACAGGTACAAGAGGATGAGAGGGTGTGTAACGATCCATGAAATC 3297
Db 3226 CTCGGTGTACAGGTACAAGAGGATGAGAGGGGATGAGAGGGGCTGCGTAACGATCCATGAAATC 3285
```

```
Qy 3298 GAGAACAAATACAGACGAACCTAAATTTAAAACTGTGAAGAGGAGGAGTGTATCCAAACG 3357
Db 3286 GAGAACAAATACAGACGAACCTAAATTTAAAACTGTGAAGAGGAGGAGTGTATCCAAACG 3345
Qy 3358 GATACAGGAACGTGTAATGATTTACTGACACCAACCAAGGTACAGCTGGATCGGACGATGCA 3405
Db 3346 GATACAGGAACGTGTAATGATTTACTGCAACCAAGGTACAGCTGGATCGGACGATGCA 3405
Qy 3406 TGTAAATTCCTGTAATGCTGGATATGAGGATGCAATATGAAGTTGATACACAGCATCTGTT 3465
Db 3406 TGTAAATTCCTGTAATGCTGGATATGAGGATGCAATATGAAGTTGATACACAGCATCTGTT 3465
Qy 3466 AATTACAAACCGACTTATGAAGAGAAACGATATACAGATGTAAGAGAGATATCATTTGT 3525
Db 3466 AATTACAAACCGACTTATGAAGAGAAACGATATACAGATGTAAGAGAGATATCATTTGT 3525
Qy 3526 GAATATGACAGAGGGTATGTAATTTACCACTACCACTACCACTGTTATATGACAAAAAGAA 3585
Db 3526 GAATATGACAGAGGGTATGTAATTTACCACTACCACTACCACTGTTATGACAAAAAGAA 3585
Qy 3586 TTAGAATACTTCCAGAAACCGATTAAGGTATGAGATTGAGATTGAGAAACCGAAGGGAAG 3645
Db 3586 TTAGAATACTTCCAGAAACCGATTAAGGTATGAGATTGAGATTGAGAAACCGAAGGGAAG 3645
Qy 3646 TTTATTGTAGACAGCGTGAATTTACTCTTATGAGGAATAG 3687
Db 3646 TTTATTGTAGACAGCGTGAATTTACTCTTATGAGGAATAG 3687

RESULT 9
AF368257
LOCUS
DEFINITION
  Bacillus thuringiensis plasmid encoded Cry1Ba (cry1Ba) gene,
  complete cds.
ACCESSION
  AF368257
VERSION
  AF368257.1 GI:14486713
KEYWORDS
  Bacillus thuringiensis
SOURCE
  Bacillus thuringiensis
ORGANISM
  Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
  cereus group.
REFERENCE
  1. (bases 1 to 3687)
  Zhang, J., Song, F. and Huang, D.
  Direct Submission
  Submitted (05-APR-2001) Biotechnology Research Group, IPP, CAAS, 2
  West Yuanmingyuan Road, Beijing, China
  Location/Qualifiers
    1..3687
    /organism="Bacillus thuringiensis"
    /mol_type="genomic DNA"
    /db_xref="taxon:1428"
    /notes="plasmid-encoded"
    1..3687
    /genes="cry1Ba"
    1..3687
    /genes="cry1Ba"
    /notes="endotoxin"
    /codon_start=1
    /transl_table=11
    /product="Cry1Ba"
    /protein_id="AAK63251.1"
    /translation="MTSNRKNENIINAVNSHAQMDLIPDARIEDSLCIAENNNIDP
    FYSATVQTGINIAGRIILGVFPAGOLASFYFLVGLWPRGRQWEIFLSEHVEOL
    INQITENARTALARLQGLGDSFRAYQQSLEDMLNDRDAPRSVLTYQVIALELDF
    LNAPELFAIRNQEVPLIMVTAQNALHLLLRDASLFGSEGLTSEIQRIYERQVER
    TRDSYDCVWEYNTGLNSLRGTNAASWVYNPFRDLTLGLDLVALFSLFYDRTYPI
    TNSAQITREVYTDATGATGVNMAWMYNNAPFSIAIAAAIRSPHLLDFLQLTIF
    SASRWNTRHMTYWRGTPIQSRPIGGGLTSTHGATNTSINPVLRFASRDFVRYTES
    YAGVLLWGLYLEPIHGVPTVRFNTPONTSDRGATANYSQPYSPGLQKSDSELP
    ETTERENYSYSHRLSHIGIILQSRVNVFVSWTHRSADRTNITGNRIQIIPWKA
    ELPPQGTVVVRGPGFTGLIRNTNCGFGPIRVTVNGPLTQRIQIRFYASTVDFF
    VSRGGTTVNNFRFLRTMNSGDELKYNFVRRAFTTPTFTTQIQDIIRTSIQGLSGNE
```

VYIDKIEIIPVATFEAYDLERAQEVNALFTNPRRLKTDVTDHYIDQVSNLVAC
 LSDEFCLDEKLEKVKYAKRLSDERNLLQDPNFTSINKOPDFISINBOSNFTS
 QSEHGWSENIITQEDNCPKENVNLTGTFNFCYPTLYXOKIGESLKYATFYOLR
 GYIEDSODILEYLRYNKAKHETLDPGTESLWPLSVESPIGRCEPNRCAPHFENFED
 LDSCRDEKCAHSHHPSLDIDVGCCTDLHENLVGVVFKIKTOEGHARLGNLEFIEE
 KPLLEALSRVKRAKKWRDKREKLOLETRVYTEAKEAVDALFVDSQYDRLOQNTNI
 GMIHAADKLVRHREAYLSLPLVPGVNAEIVFEELGHIITATSLYDARNVKNQDFN
 NGLTCWNVKGVHDVQOQSHRSDLVPEWEAEVSOAVRCPGCGVILKVTAYKEGYGBG
 NLTLEINNTDELKFNKREEEVYPTDTGTCTNDYTAHQGTAGCADACSNRAGYEDA
 YEVDTTASVNYKPYESETYDVRDNHCEYDRGVYVNPVPPAGVTVTKELLEYFETDT
 VWIEIGETEGKFIIVDSVELLMEE"

ORIGIN

Query Match 85.0%; Score 3135.2; DB 1; Length 3687;
 Best Local Similarity 91.1%; Pred. No. 0;
 Matches 3374; Conservative 0; Mismatches 298; Indels 30; Gaps 3;

Qy	1	TTGACTTCAAAATAGAAAATGAGAAATGAGAAATTAATAATGCTTTATCGATTTCCAGCTGTA	60
Db	1	TTGACTTCAAAATAGAAAATGAGAAATGAGAAATTAATAATGCTTTATCGATTTCCAGCTGTA	45
Qy	61	TCGAATCATTCACACAAATGATCTATCACCAGATGCTCGTATTGAGGATTTCTTGTGT	120
Db	46	TCGAATCATTCACACAAATGATCTATCACCAGATGCTCGTATTGAGGATTTCTTGTGT	105
Qy	121	ATAGCCGAGGGGAATAATCAATCCACTTGTAGCGCATCAACAGATCCAAACGGGTATT	180
Db	106	ATAGCCGAGGGGAATAATCAATCCACTTGTAGCGCATCAACAGATCCAAACGGGTATT	165
Qy	181	AACATAGCTGTAGAAATCTAGTGTATTAGCGGTAACGGTTGTCGACAAATAGCTAGT	240
Db	166	AACATAGCTGTAGAAATCTAGTGTATTAGCGGTAACGGTTGTCGACAAATAGCTAGT	225
Qy	241	TTTTATAGTTTTCTTGTGTGTAATTATGCGCGCGCGAGATCAGTGGGAAATTTTC	300
Db	226	TTTTATAGTTTTCTTGTGTGTAATTATGCGCGCGCGAGATCAGTGGGAAATTTTC	285
Qy	301	CTAGAACATGTGCAACAACTTATAAATCAACAAATAACAGAAATGCTAGGAATACGGCA	360
Db	286	CTAGAACATGTGCAACAACTTATAAATCAACAAATAACAGAAATGCTAGGAATACGGCT	345
Qy	361	CTTGCTCGATTAAAGTTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT	420
Db	346	CTTGCTCGATTAAAGTTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT	405
Qy	421	TGGCTAGAAAACCTGTATGATGCAAGAACGAGAGTGTCTTTATACCCCAATATAGCC	480
Db	406	TGGCTAGAAAACCTGTATGATGCAAGAACGAGAGTGTCTTTATACCCCAATATAGCT	465
Qy	481	TTAGAACTTGATTTTCTTAATGCGATCGCGCTTTTCGCAATTAGAAACCAAGAGTTCCA	540
Db	466	TTAGAACTTGATTTTCTTAATGCGATCGCGCTTTTCGCAATTAGAAACCAAGAGTTCCA	525
Qy	541	TTATTATGTTATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGCTCT	600
Db	526	TTATTATGTTATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGCTCT	585
Qy	601	CTTTTGTGTAGTAAATTTGGGCTTTACATCGCAGAAATTCACGTTATTATGAGCGCAA	660
Db	586	CTTTTGTGTAGTAAATTTGGGCTTTACATCGCAGAAATTCACGTTATTATGAGCGCAA	645
Qy	661	GTGGAAACAAACGAGAGATTATTCGACTATTGGGTAGAAATGGTATATACAGCTCTAAAT	720
Db	646	GTGGAAACAAACGAGAGATTATTCGACTATTGGGTAGAAATGGTATATACAGCTCTAAAT	705
Qy	721	AGCTTGAGGGGCAAAATGCGCAAGTTGGTGGGTATTATCAATTCGATAGATCTA	780
Db	706	AGCTTGAGGGGCAAAATGCGCAAGTTGGTGGGTATTATCAATTCGATAGATCTA	765
Qy	781	ACGTTAGGGCTATTAGATCTAGTGGCACTATTCCCAAGCTATGACACTCGCACTTATCCA	840
Db	766	ACGTTAGGGCTATTAGATCTAGTGGCACTATTCCCAAGCTATGACACTCGCACTTATCCA	825

Qy	841	ATAAATACAGTGTCTCAGTTTAAACAGGGAAGTTTATACAGACGCAATTTGAGCAACAGGG	900
Db	826	ATAAATACAGTGTCTCAGTTTAAACAGGGAAGTTTATACAGACGCAATTTGAGCAACAGGG	885
Qy	901	GTAATAATGCAAGTATGAATTTGGTATAATAATGCAACCTTCGTTTTCCGCTATAGAG	960
Db	886	GTAATAATGCAAGTATGAATTTGGTATAATAATGCAACCTTCGTTTTCCGCTATAGAG	945
Qy	961	ACTGCGGTATTCGAAGCCCGCATCTACTTGAATTTCTTAGAACAACTTACAAATTTTACG	1020
Db	946	GCTGCGGCTATTCGAAGCCCGCATCTACTTGAATTTCTTAGAACAACTTACAAATTTTACG	1005
Qy	1021	ACTTTCATCAGATGGAGTGTCTACTAGCATATGACTTACTTGGCGGGGCGACAAATTTCAA	1080
Db	1006	GCTTTCATCAGATGGAGTGTCTACTAGCATATGACTTATTGGCGGGGCGACAAATTTCAA	1065
Qy	1081	TCTCGGCGCAATAGGAGCGGATTTAAATACCTCAACGCATGGGTCTACCAATATCTTCTATT	1140
Db	1066	TCTCGGCGCAATAGGAGCGGATTTAAATACCTCAACGCATGGGTCTACCAATATCTTCTATT	1125
Qy	1141	AATCCTGTAGATATCATTTCTCTCTGAGAGCTATATTGGACTGAATCATATGACGGA	1200
Db	1126	AATCCTGTAACTTACCGTTTCGCAATCTCGAGACGTTTATAGGACTGAATCATATGACGGA	1185
Qy	1201	GTGCTTCTATGGGGAATTTTACCTTTGAACCTTATTCTGTTGCTCCCTACTGTTAGATTAAT	1260
Db	1186	GTGCTTCTATGGGGAATTTTACCTTTGAACCTTATTCTGTTGCTCCCTACTGTTAGATTAAT	1245
Qy	1261	TTTAGGAACCTCTCAGAAATACCTTTTGAAGAGGTACTGTCTAACTATATAGTCAACCTATGAG	1320
Db	1246	TTTAGGAACCTCTCAGAAATACCTTTTGAAGAGGTACTGTCTAACTATATAGTCAACCTATGAG	1305
Qy	1321	TCACCTGGGCTTCAATTTAAAGATTCAGAACTTGAATTTACCACAGAAACAAACAGAACGA	1380
Db	1306	TCACCTGGGCTTCAATTTAAAGATTCAGAACTTGAATTTACCACAGAAACAAACAGAACGA	1365
Qy	1381	CCAAATTTATGAATCATATAGTCTATAGTGTATCTCAGATAGGGCTCAATTTCACAATCTAGG	1440
Db	1366	CCAAATTTATGAATCATATAGTCTATAGTGTATCTCAGATAGGGCTCAATTTCACAATCTAGG	1425
Qy	1441	GTGCATGTACAGTATATTCTTGGAGCGACCGTAGTGTGCGAGATCGTACAAATACCATTAGT	1500
Db	1426	GTGAATGTACCGGTATATTCTTGGAGCGACCGTAGTGTGCGAGATCGTACAAATACCATTAGT	1485
Qy	1501	TCAGATGAGCATACAAATACCATTTGTTAAATCATTTCAACCTTAAATTCAGGTACTCTCT	1560
Db	1486	CCAAATAGAAATCACCAATTCCTATGTTAAAGCATCGAACTTCTCCTCAAGGTACCACT	1545
Qy	1561	GTAGTCAGTGGCCAGGATTTTACAGGAGGGATATTAATCCGAACTAACGTTAATGGTAGT	1620
Db	1546	GTGTTTAGAGGACCAAGATTTTACTGTTGGGATATTCTTTCGAGAGACGAATACCTGGTGA	1605
Qy	1621	GTAATAAGTATGGGTCTTAATTTTAAATATACATCATTTACAGCGGTATCGGTGAGAGTT	1680
Db	1606	TTTGGACCGATAAGAGTAACTGTTTAACCGGACCATTAACACAAAGATATCGTATAGGATTC	1665
Qy	1681	CGTTATGCTGCTTCTCAACCAATGGTCTGAGGGTAACTGTCGAGGAGGAGTACTACTTTT	1740
Db	1666	CGCTATGCTTCAACTGTAGATTGTTTCTTCTTGTATACGTGGAGGTACTACTGTAAAT	1725
Qy	1741	GATCAAGGATTCCTTAGTACTATAGTGCATAATGAGTCTTTGCACTCTCAATCAITTTAGA	1800
Db	1726	AATTTTAGATTTCTACGTACAAATGAACAGTGGAGACGAATTAATAACGAAATTTTGTG	1785
Qy	1801	TTTGCAAGATTTTCTGTAGTATTAGTGCATCTGGCAGTCAA---ACTGCTGGAATAAGT	1857
Db	1786	AGACGTGCTTTTACTACACCTTTTACTTTTACACAAATTTCAAGATATAATTCGAACGTCT	1845
Qy	1858	ATAAGTAATAATGACAGTACAAACGTTTCTTGTATAAATAATTTCAATTCATTTCAATTT	1917
Db	1846	ATTCAAGGCTTTAGTGGAAATGGGGAAGTGTATATAGATATAAATTTGAATTTTCCAGTT	1905
Qy	1918	ACTGCAACCTTCGAGCAGAAATACGATTTTAGAAAGGGCGCAAGCGGTGAATGCTCTG	1977

1906 ACTGCAACCTTCGAGCAGAAATGATTTAGAAAGCGCAGAGCGGTGAATGCTCTG 1965
1978 TTTTACTAATACGAATCCAAGAAGATTGAAACAGATGTGACAGATTATCATATTGATCAA 2037
1966 TTTTACTAATACGAATCCAAGAAGATTGAAACAGATGTGACAGATTATCATATTGATCAA 2025
2038 GTATCCAAATTTAGTGGCGTGTATTCGGATGAATTCCTGTAGATGAAAGAGAGAAATTA 2097
2026 GTATCCAAATTTAGTGGCGTGTATTCGGATGAATTCCTGTAGATGAAAGAGAGAAATTA 2085
2098 CTTGAGAAAGTGAATATGCGAACGACTCAGTGCATGAAAGAACTTACTCCAGATCCA 2157
2086 CTTGAGAAAGTGAATATGCGAACGACTCAGTGCATGAAAGAACTTACTCCAGATCCA 2145
2158 AACTTCACATCCATCAATTAAGCAACCAAGACTTCATATCTACTAATGAGCAATCGAATTTTC 2217
2146 AACTTCACATCCATCAATTAAGCAACCAAGACTTCATATCTACTAATGAGCAATCGAATTTTC 2205
2218 ACATCTATCCATGAACAACTCTGAACATGGAATGGTGGGGAAGTGAGAACATTAACAATCCAG 2277
2206 ACATCTATCCATGAACAACTCTGAACATGGAATGGTGGGGAAGTGAGAACATTAACAATCCAG 2265
2278 GAAGGAATACCGTATTTAAAGAAATTAAGTACACTACCGGGAATTTTAATGAGTGT 2337
2266 GAAGGAATACCGTATTTAAAGAAATTAAGTACACTACCGGGAATTTTAATGAGTGT 2325
2338 TATCCGACGTATTTATATCAAAAATAGGAGAGTCCGGAATTTAAAGCTTATATCTCGCTAC 2397
2326 TATCCGACGTATTTATATCAAAAATAGGAGAGTCCGGAATTTAAAGCTTATATCTCGCTAC 2385
2398 CAATTAAGAGGTATATTTGAAGATGATCAAGATTTAGAGATATATTTGATTCGTTATAAT 2457
2386 CAATTAAGAGGTATATTTGAAGATGATCAAGATTTAGAGATATATTTGATTCGTTATAAT 2445
2458 GCGNAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGTCGATGGCGCTTTCAGTTGAA 2517
2446 GCGNAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGTCGATGGCGCTTTCAGTTGAA 2505
2518 AGCCCAATCGAAGTCCGAGAACCCGAATCGATCGCACACACATTTTGAATGGAATCCCT 2577
2506 AGCCCAATCGAAGTCCGAGAACCCGAATCGATCGCACACACATTTTGAATGGAATCCCT 2565
2578 GATCTAGATTTCTCCTCAGAGATGGAGAAAAATGTCCGCATCATTTCCCATCATTTCTCT 2637
2566 GATCTAGATTTCTCCTCAGAGATGGAGAAAAATGTCCGCATCATTTCCCATCATTTCTCT 2625
2638 TTGGATATTTGATTTGATGACAGACTTGCATGAGAAATCTAGCGGTGTGGTGTATTC 2697
2626 TTGGATATTTGATTTGATGACAGACTTGCATGAGAAATCTAGCGGTGTGGTGTATTC 2685
2698 AAGATTAAAGCGCAGGAAGTTCATGCAAGACTAGGGAATCTGGAATTTATTGAAGAGAAA 2757
2686 AAGATTAAAGCGCAGGAAGTTCATGCAAGACTAGGGAATCTGGAATTTATTGAAGAGAAA 2745
2758 CCATTTATTAGGAGAACACTGTCTCGTGTGAAGAGAGCAGAGAAAAAATGGAGACAAA 2817
2746 CCATTTATTAGGAGAACACTGTCTCGTGTGAAGAGGCGCAGAGAAAAAATGGAGACAAA 2805
2818 CGTGAAGAACTACAAATTTGAAAAAAGAGTATATACAGAGGCAAAAGAGCTGTGGAT 2877
2806 CGTGAAGAACTACAAATTTGAAAAAAGAGTATATACAGAGGCAAAAGAGCTGTGGAT 2865
2878 GCTTTATTGTAGATTTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATGATT 2937
2866 GCTTTATTGTAGATTTCTCAATATAATAGATTACAAGCGGATACAAACATCGGCATGATT 2925
2938 CATCGGCAGATAAACTTTGTTTCATCGAATTCGAGAGGCTTATCTGTGAGAAATTTATCTGTT 2997
2926 CATCGGCAGATAAACTTTGTTTCATCGAATTCGAGAGGCTTATCTTTGAGAAATTTACCTGTT 2985
2998 ATCCCGGCTGTAATGCGGAATTTTGAAGAAATTTAGAGGTCGCATTTATCATCTGCAATC 3057

2986 ATCCAGGTGTAATGCGGAATTTTGAAGAAATTAGAAAGTCACATTATCATCTGCAATC 3045
3058 TCCTATACGATCGAGAAATGTCGTTAAAAATGGTGATTTTAATAATGATTAGATGAC 3117
3046 TCCTTATACGATCGAGAAATGTCGTTAAAAATGGTGATTTTAATAATGATTAGATGAC 3105
3118 TGGAAATGTAAGGGGATGTAGATGTACACAGAGCCATCACCGTTCCTGCTGTTATC 3177
3106 TGGAAATGTAAGGGGATGTAGATGTACACAGAGCCATCATCGTTCCTGCTGTTATC 3165
3178 CCAGAAATGGAAGCAGAAAGTGTCAAGCAGTTCGCGTCTGTCCGGGCGTGGCTATATC 3237
3166 CCAGAAATGGAAGCAGAAAGTGTCAAGCAGTTCGCGTCTGTCCGGGCGTGGCTATATC 3225
3238 CTCGCTGTCAAGCGTCAAAAGAGGATATGGAAGGGTGTGTGTAACGATCCATGAAATC 3297
3226 CTTTCGCTGTCACGCTACAAAGAGGGATATGGAAGGGCTGCGTAAACGATCCATGAAATC 3285
3298 GAGAACATACAGACGAACTTAAATTTTAAAACTGTGAAGAGGAGGATGTATCCAAAG 3357
3286 GAGAACATACAGACGAACTTAAATTTTAAAACTGTGAAGAGGAGGATGTATCCAAAG 3345
3358 GATACAGGAACGCTGTAATGATTTACTGTCACACCAAGGTATACAGC-----AGTA 3405
3346 GATACAGGAACGCTGTAATGATTTACTGTCACACCAAGGTATACAGC-----AGTA 3405
3406 TGTAAATTCCTGATGCTGGATATGAGGATGCAATATGAAGTTGATATACAGCATCTGTT 3465
3406 TGTAAATTCCTGATGCTGGATATGAGGATGCAATATGAAGTTGATATACAGCATCTGTT 3465
3466 AATTACAAACCGCTTATGAAGAAACGATATACAGATGTACGAGAGATATCATCTGT 3525
3466 AATTACAAACCGCTTATGAAGAAACGATATACAGATGTACGAGAGATATCATCTGT 3525
3526 GAATATGACAGAGGATGTGAATTTATCCACCTACCGCTGTTATATGACAAAAAGAA 3585
3526 GAATATGACAGAGGATGTGAATTTATCCACCTACCGCTGTTATATGACAAAAAGAA 3585
3586 TTAGAATCTTCCAGAAACCGATTAAGGTATGAGATGAGATGGAAGAACGAGAGGAAG 3645
3586 TTAGAATCTTCCAGAAACCGATTAAGGTATGAGATGAGATGGAAGAACGAGAGGAAG 3645
3646 TTTATTGTACAGCGTGGAAATTTACTCTTATGGAGGAATAG 3687
3646 TTTATTGTAGTAGCGTGGAAATTTACTCTCATGGAAGATAG 3687

RESULT 10
BTRV1BA2
LOCUS
DEFINITION
B.thuringiensis cry1Ba2 gene.
ACCESSION
X95704
VERSION
X95704.1 GI:1200001
KEYWORDS
cry1Ba2 gene.
SOURCE
Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE
1
AUTHORS
Soetaert, P.
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 4074)
AUTHORS
Soetaert, P.
TITLE
Direct Submission
JOURNAL
Submitted (15-FEB-1996) P. Soetaert, Plant Genetic Systems, R&D, J. Plateausstraat 22, Gent, B 9000, Belgium
FEATURES
source
1. .4074
Location/Qualifiers
/organism="Bacillus thuringiensis"
/mol_type="genomic DNA"
/strain="HD-110"
/variety="entomocidus"
/db_xref="taxon:1428"

BTRV1BA2 4074 bp DNA linear BCT 21-FEB-1996
B.thuringiensis cry1Ba2 gene.
X95704
X95704.1 GI:1200001
cry1Ba2 gene.
Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
1
Soetaert, P.
Unpublished
2 (bases 1 to 4074)
Soetaert, P.
Direct Submission
Submitted (15-FEB-1996) P. Soetaert, Plant Genetic Systems, R&D, J. Plateausstraat 22, Gent, B 9000, Belgium
1. .4074
Location/Qualifiers
/organism="Bacillus thuringiensis"
/mol_type="genomic DNA"
/strain="HD-110"
/variety="entomocidus"
/db_xref="taxon:1428"

```

RBS      174. .177
gene      /gene=.3872
CDS       /gene=.cry1Ba2
          /codon_start=1
          /transl_table=11
          /protein_id="CAA65003.1"
          /db_xref="GI:1200002"
          /db_xref="GOA:P05517"
          /db_xref="UniProt/Swiss-Prot:P05517"
          /translation="MTSNRKNENIINAVSNHSAQMDLLPDARIEDSLCIAEAGNNIDP
          FVSASTVQTGINAGRIQVLGVFPAGQASFYSLVGEIWPGRGDQWEIFLEHVEOL
          INQOITENARTALRQLQGLDSFRAYQOSLELDWLENKRDARTSRVLIQVIALELDF
          LNAFLPAIRNOEVLWVYAOAANLHLLLRDLASLFGSEBGLTSQSIQRYVEROER
          TRDYSVCVEVNTGLSLRGTNAASVWVYNQPERDLTGLVLDLVALFPSYDRTYFI
          NTSAQLREVTYDAIGATGVNMAWNNNAPSFAIEAAHKSPLLDFLEQITIF
          SASGRNRMHTWRIGHTQSRIPIGGLNLTSHGATNTSINPVTLRFASRDVYRTES
          YAGVLLWGIYLEPITHGVPTVRFPNTSDRGATANYSPQSPYSGQLKDSLETLP
          ETTERPNYESHRLSHIGILQSRVNVVSVWTHRSADRNTNIPNRITQIPWVKAS
          ELPOGTTVVRPGFTGDIILRRTTGGVGRIRVTNGLPTQVRIGRYASTVDGDFP
          VSRGTTVNNRPLRTWNSGDELKYGNFVRAFTTPTFTQIODIISTSLQGLSGNGE
          VYIDKIELIPWTAFEAEDYLERAEVNAIPLFNTNPRRLKTDVTDHIDQVSNLVC
          LSDFCDEKRELLEKVKAKRLSDERNLLQDPNFTSINKOPDFISNEOSNFTSIHE
          QSEHWGSENIITIEQDNVFNENYVTLPGTFNECYPTIYLYKIGESLKAITYRQLR
          GYIEDSODLEIYLRYNKHEHETLDVPGTESLWPLSVESPIRCGEPNRCAPHFEMNP
          LDSCRDEKCAHSHHFLSDIDVGCTDLHENLGVWVFKITQEBGHARLGNLEFIE
          KPLLEALSRLVRKAEKWRDKREKLOLETKRVVTEAKEAVALFVDSQYDRLOADINI
          GMUHAADKLVRHIREAVLSELPLVPGVNAEIPFELEGHIIITAIISLDARNVXNGDN
          NGLTCWNKGVHDVQQSHHRSDDLVIPEWAEVSAVRVCPGCGTILKRTAYKSGYGG
          YEVDTHTASVNYKPTIEETVTDVDRDNHCEYDRGVNYPVPVAGYVTKLEYFPETDT
          VWIEIGTEGKFIQVDSVELLME"
          3958. .3981
          terminator
          ORIGIN

Query Match      85.0%; Score 3133.6; DB 1; Length 4074;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 3373; Conservative 0; Mismatches 299; Indels 30; Gaps 3;

QY      1  TTGACTTCAATAGAAAATGAGAATGAATATAATGCTTTATCGATTCCAGCTGTA 60
DB      186 TTGACTTCAATAGAAAATGAGAATGAATATAATGCTTTATCGATTCCAGCTGTA 230

QY      61  TCGAATCATTCACACAAATGGATCTATCACAGATGCTCGTATGTAGGAGTTCCTTGTGT 120
DB      231 TCGAATCATTCGCGACAAATGGATCTATTACCAGATGCTCGTATGTAGGATAGCTTGT 290

QY      121 ATAGCCGAGGGGATATATCAATCCACTGTTAGCGCATCAACAGTCCAAACGGGTATT 180
DB      291 ATAGCCGAGGGGAAACAATATGTATCCCAATTTGTTAGCGCATCAACAGTCCAAACGGGTATT 350

QY      181 AACATAGCTGTAGAACTACTAGGTGTATTAGCGGTACCGTTTGTGTCGACAAATAGCTAGT 240
DB      351 AACATAGCTGTAGAACTACTAGGTGTATTAGCGGTACCGTTTGTGTCGACAACTAGCTAGT 410

QY      241 TTTTATAGTTTTCTGTGTGTGAATATTAGCCCGCGGAGAGATCAGTCGGGAAATTTTC 300
DB      411 TTTTATAGTTTTCTGTGTGTGAATATTAGCCCGCGGAGAGATCAGTCGGGAAATTTTC 470

QY      301 CTAGAACATCTCGAACAACTTATAATCAACAAATAACAGAAATGCTAGGAATACGGCA 360
DB      471 CTAGAACATCTCGAACAACTTATAATCAACAAATAACAGAAATGCTAGGAATACGGCT 530

QY      361 CTGTCTCGATTACAAGGTTTAGGAGATTCCTTTAGAGCCCTATCAACAGTCACCTTGAAGAT 420
DB      531 CTGTCTCGATTACAAGGTTTAGGAGATTCCTTTAGAGCCCTATCAACAGTCACCTTGAAGAT 590

QY      421 TGGCTAGAAAACCGGTGATGATGCAAGAAACGAGAAAGTGTCTTTATATCCCAATATATAGCC 480
DB      591 TGGCTAGAAAACCGGTGATGATGCAAGAAACGAGAAAGTGTCTTTATATCCCAATATATAGCT 650

QY      481 TTAGAACTTGATTTCTTTAATGCGATGCCCTTTTCGCAATTAGAACCAAGAGTTCCA 540

```

```

DB      651 TTAGAACTTGATTTCTTAATGCGATGCCCTTTTCGCAATTAGAACCAAGAGTTCCA 710
QY      541 TTATTAAATGATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGCTCTCT 600
DB      711 TTATTGATGATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGCTCTCT 770
QY      601 CTTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAATTCGATTTATTATTAGGCGCAA 660
DB      771 CTTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAATTCGATTTATTATTAGGCGCAA 830
QY      661 GTGGAACAAACGAGAGAGATTTATTCGCACTATTGCTGAGATGCTGATATATACAGGTCTAAAT 720
DB      831 GTGGAACGAAACGAGAGAGATTTATTCGCACTATTGCTGAGATGCTGATATATACAGGTCTAAAT 890
QY      721 AGCTTGAGAGGGAACAATGCGCAAGTTGGTGGCTGTTTAAATCAATTCCTGAGAGATCTA 780
DB      891 AGCTTGAGAGGGAACAATGCGCAAGTTGGTGGCTGTTTAAATCAATTCCTGAGAGATCTA 950
QY      781 ACCTTAGGGGTATTAGATCTAGTGGCACTATTCGCAAGCTATGACACTCGCAGCTTATCCA 840
DB      951 ACCTTAGGAGTATTAGATCTAGTGGCACTATTCGCAAGCTATGACACTCGCAGCTTATCCA 1010
QY      841 ATAAATACGAGTCTCAGTTTAAACAGGGAAGTTTATATACAGACCAATTTGAGGCAACAGGG 900
DB      1011 ATAAATACGAGTCTCAGTTTAAACAGGGAAGTTTATATACAGACCAATTTGAGGCAACAGGG 1070
QY      901 GTAAATATGCAAGTATGAATTCGTTTAAATTAATTAATGACCTTCGTTTCCGCTATAGAG 960
DB      1071 GTAAATATGCAAGTATGAATTCGTTTAAATTAATTAATGACCTTCGTTTCCGCTATAGAG 1130
QY      961 ACTGCGGTATTCCGAGCCGCACTACTTCTGATTTCTTAGAACAACTTACAAATTTTATAGC 1020
DB      1131 GCTGCGGTATTCCGAGCCGCACTACTTCTGATTTCTTAGAACAACTTACAAATTTTATAGC 1190
QY      1021 ACTTCATACAGTGGAGTGTCTACTAGGCATATGCACTTACTGCGGGGGGACACAAATTCAA 1080
DB      1191 GCTTCATACAGTGGAGTGTCTACTAGGCATATGCACTTACTGCGGGGGGACACAAATTCAA 1250
QY      1081 TCTCGGCCAATAGGAGCGGATTAATACCTCAACGATGCGGTCTACCAATCTCTTATT 1140
DB      1251 TCTCGGCCAATAGGAGCGGATTAATACCTCAACGATGCGGTCTACCAATCTCTTATT 1310
QY      1141 AATCCTGTGAAGTATCATTTCTCTCTCGAGAGCTATATTGGAAGTCAATCATATGAGGAG 1200
DB      1311 AATCCTGTGAAGTATCATTTCTGAGAGCTTTATAGGAGTCAATCATATGAGGAG 1370
QY      1201 GTGCTTCTATGGGAATTTTACCTTGAACCTTATTCATGTTGTCCTACTGTTAGATTTAAT 1260
DB      1371 GTGCTTCTATGGGAATTTTACCTTGAACCTTATTCATGTTGTCCTACTGTTAGATTTAAT 1430
QY      1261 TTTTAGGAACCTTCAGAAATTAATTTTGAAGAGGTACTGCTAACTATATAGTCAACCTATGAG 1320
DB      1431 TTTTAGGAACCTTCAGAAATTAATTTTGAAGAGGTACTGCTAACTATATAGTCAACCTATGAG 1490
QY      1321 TCACCTGGGCTTCAATTAAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAA 1380
DB      1491 TCACCTGGGCTTCAATTAAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAA 1550
QY      1381 CCNAATTAATGAATCATATATAGTCAATGTTTATCTCATATAGGCTCATTTTCAATCTAGG 1440
DB      1551 CCNAATTAATGAATCATATATAGTCAATGTTTATCTCATATAGGCTCATTTTCAATCTAGG 1610
QY      1441 GTGCAATGACCAAGTATATTTCTTGGAGCGCAACCGTAGTGCAGATCGTACAAATACCAATAGT 1500
DB      1611 GTGAATGACCGGTATATTTCTTGGAGCGCATCGTAGTGCAGATCGTACAAATACCAATAGT 1670
QY      1501 TCAGATAGCAATACCAAAATACAAATGATGTTGTTAAATCATTTCAACTTTAGGAGTCTCT 1560
DB      1671 CCNAATAGAAATACCCAAATCCAAATGTTGTTAAAGCATCCGAACTTCTCCTCAAGGTACCACT 1730
QY      1561 GTAGTCAGTGGCGGAGGATTTTACAGGAGGAGTATTAATCCGAACTAAACGTTAATGTTAGT 1620

```


Db 1731 GTTGTAGGACGAGGATTTACTGGTGGGATATCTTCGAAAGCAAGATCTGGTGA 1790
 Qy 1621 GTACTAAGTAGTGGTCTTAATTTAAATAATACATCAATACAGCGGTATCGCGTGAGAGTT 1680
 Db 1791 TTTGGACCGATAAGAGTAACTGTCTAAACGACCAATTAACACAAAGATATCGTATAGGATTC 1850
 Qy 1681 CGTTATGCTGCTCTCAAAACAATCGTCTGAGGCTAACTGTCGAGGAGTACTACTTTT 1740
 Db 1851 CGCTATGCTTCACTGATGATTTTGAATTTCTTTGATACAGTGGAGGTACTACTGTAAAT 1910
 Qy 1741 GATCAAGGATTCCTAGTACTATGAGTGCATAATGAGTCTTTGACATCTCAATCAATTTAGA 1800
 Db 1911 AATTTTAGATTCCTACGTACATGCAAGTGGAGACGCAATTAATAATACGGAAATTTTGTG 1970
 Qy 1801 TTTCCGAAATTTCTGTAGTATTTAGTGCATCTGGCAGTCAA---ACTGCTGGAATAAGT 1857
 Db 1971 AGACGTGCTTTTACTACACCTTTTACTTTTACACAAATTCAGATATAAATTCGAACGCTCT 2030
 Qy 1858 ATAAGTAAATGACAGTACGAGTACAAACGTTTCACTTTTGTATTAATAATTTGAATTTCCAAT 1917
 Db 2031 ATTCAGGCTTGTAGTGAATGGGAAAGTGTATATGATATAAATTTGAAATTTATTTCCAGTT 2090
 Qy 1918 ACTGCAACCTTCGAAGCAGATACGATTTAGAAAGGCGCAAGAGCGGTGAATGCTCTG 1977
 Db 2091 ACTGCAACCTTCGAAGCAGATATGATTTAGAAAGGCGCAAGAGCGGTGAATGCTCTG 2150
 Qy 1978 TTTACTAATACGAATCCAAAGAAGTTGAAACAGATGTGACAGATTTATCATATTTGATCAA 2037
 Db 2151 TTTACTAATACGAATCCAAAGAAGTTGAAACAGATGTGACAGATTTATCATATTTGATCAA 2210
 Qy 2038 GTATCCAAATTTAGTGGGCTGTTATCGGATGAATTTCTGTTAGATGAATAAGAGAAATTA 2097
 Db 2211 GTATCCAAATTTAGTGGGCTGTTATCGGATGAATTTCTGTTAGATGAATAAGAGAAATTA 2270
 Qy 2098 CTTGAGAAAGTGAATATCGAAACGACTCAGTGCATGAAGAACTTACTCCAGATCCA 2157
 Db 2271 CTTGAGAAAGTGAATATCGAAACGACTCAGTGCATGAAGAACTTACTCCAGATCCA 2330
 Qy 2158 AACTTCATCCATCAATTAAGAACACAGACTTCATATCTACTAATAGCAATCGAATTTTC 2217
 Db 2331 AACTTCATCCATCAATTAAGAACACAGACTTCATATCTACTAATAGCAATCGAATTTTC 2390
 Qy 2218 ACATCTATCAATGAACATCTGAACATGATGATGGGGAAGTGAGAACATTAACATCCAG 2277
 Db 2391 ACATCTATCCATGAACATCTGAACATGATGATGGGGAAGTGAGAACATTAACATCCAG 2450
 Qy 2278 GAAGGAATGACGTTATTTAAGGAATTAAGTGCACACTACCGGAGCTTTTATGAGTGT 2337
 Db 2451 GAAGGAATGACGTTATTTAAGGAATTAAGTGCACACTACCGGAGCTTTTATGAGTGT 2510
 Qy 2338 TATCCGACGTTATTTATCAAAAATAGGAGAGTCGGAATTTAAAGCTTATACTCGCTAC 2397
 Db 2511 TATCCGACGTTATTTATCAAAAATAGGAGAGTCGGAATTTAAAGCTTATACTCGCTAC 2570
 Qy 2398 CAATTAAGAGGTATATGAAATGATGCAAGATTTAGAGATATATTTGATTCGTTATAAT 2457
 Db 2571 CAATTAAGAGGTATATGAAATGATGCAAGATTTAGAGATATATTTGATTCGTTATAAT 2630
 Qy 2458 GCGAAACATGAACATTTGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTCAGTTGAA 2517
 Db 2631 GCGAAACATGAACATTTGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTCAGTTGAA 2690
 Qy 2518 AGCCCAATCGAAGTCCGAGAACCGAATCGATGCGCACACATTTTGAATGGAATCCCT 2577
 Db 2691 AGCCCAATCGAAGTCCGAGAACCGAATCGATGCGCACACATTTTGAATGGAATCCCT 2750
 Qy 2578 GATCTAGATTTGCTCTGACAGATGAGAAAAATGTGCGCATCATTTCCCATCAFTTCTCT 2637
 Db 2751 GATCTAGATTTGCTCTGACAGATGAGAAAAATGTGCGCATCATTTCCCATCAFTTCTCT 2810
 Qy 2638 TTGATATTTGATATTTGATGACAGACTTGTGATGAGAAATCTAGCGGTGCGGTGATATTC 2697
 Db 2811 TTGATATTTGATATTTGATGACAGACTTGTGATGAGAAATCTAGCGGTGCGGTGATATTC 2870

RESULT 11
 142103
 LOCUS

142103

4074 bp

DNA

linear

PAT 07-OCT-1997

Qy 2698 AAGATTAAAGACGCAAGGAAGGTCTAGGAATCTGGAAATTTATTTGAAGAGAAA 2757
 Db 2871 AAGATTAAAGACGCAAGGAAGGTCTAGGAATCTGGAAATTTATTTGAAGAGAAA 2930
 Qy 2758 CCATTATTAGGAGAACACATGCTCTCGTGTGAAGAGACAGAGAAATAATGGAGAGACAAA 2817
 Db 2931 CCATTATTAGGAGAACACATGCTCTCGTGTGAAGAGGCGAGAGAAAAATGGAGAGACAAA 2990
 Qy 2818 CGTGAAAACTCAATTTGAAAAACAAACGAGTATATACAGAGGCAAAAGAGCTGTGGAT 2877
 Db 2991 CGTGAAAACTCAATTTGAAAAACAAACGAGTATATACAGAGGCAAAAGAGCTGTGGAT 3050
 Qy 2878 GCATTATTGTAGATTTCTCAATATATATAGATTACAAAGGAGATACAAACATTGGCATGATT 2937
 Db 3051 GCATTATTGTAGATTTCTCAATATATATAGATTACAAAGGAGATACAAACATTGGCATGATT 3110
 Qy 2938 CATGCGGACATAAACTTTGTTTCGAAATTCGAGAGGCTTATCTGTGAGAAATTTATCTGTT 2997
 Db 3111 CATGCGGACATAAACTTTGTTTCGAAATTCGAGAGGCTTATCTTTTCAAGAAATTTACCTGTT 3170
 Qy 2998 ATCCCGGCTGTAATTCGGAATTTTGAAGAAATTTAGAAAGTCCGATTTACCTGCAATC 3057
 Db 3171 ATCCAGGTGTAATTCGGAATTTTGAAGAAATTTAGAAAGTCCGATTTACCTGCAATC 3230
 Qy 3058 TCCTTATACGATCGAGAAATGTCGTTTAAATAATGGTGATTTTAAATAGGATTAGCATGC 3117
 Db 3231 TCCTTATACGATCGAGAAATGTCGTTTAAATAATGGTGATTTTAAATAGGATTAACTGT 3290
 Qy 3118 TGGAAATGAAGAGGCAATGTAGATGTACACAGAGCCATCACGTTCTGCTGCTTTGTTATC 3177
 Db 3291 TGGAAATGAAGAGGCAATGTAGATGTACACAGAGCCATCATCGTTCTGACCTTTGTTATC 3350
 Qy 3178 CCAGATGGGAAGCAGAAAGTGTCAAGCAGTTTCGCGTCTGTCCGGGGCGTGGCTATATC 3237
 Db 3351 CCAGATGGGAAGCAGAAAGTGTCAAGCAGTTTCGCGTCTGTCCGGGGCGTGGCTATATC 3410
 Qy 3238 CTCCTGTCTACAGCGTACAAAGAGGGATGAGAGAGGTTGTGTAAACGATCCCATGAAATC 3297
 Db 3411 CTCCTGTCTACAGCGTACAAAGAGGGATGAGAGAGGTTGTGTAAACGATCCCATGAAATC 3470
 Qy 3298 GAGAACATACAGACGAACTTAAATTTAAATACTGTGAAGAGAGGAAAGTGTATCCAAAG 3357
 Db 3471 GAGAACATACAGACGAACTTAAATTTAAATACTGTGAAGAGAGGAAAGTGTATCCAAAG 3530
 Qy 3358 GATACAGGAACGTTAATGATTTATCTGACACCAAGGTACAGC-----AGTA 3405
 Db 3531 GATACAGGAACGTTAATGATTTATCTGACACCAAGGTACAGCTGGATGCGCAGATGCA 3590
 Qy 3406 TGTAATTTCCCGTAAATGCTGATATGAGGATGCAATGAGTTGATGATACTACAGCATCTGTT 3465
 Db 3591 TGTAATTTCCCGTAAATGCTGATATGAGGATGCAATGAGTTGATGATACTACAGCATCTGTT 3650
 Qy 3466 AATTACAAACCGATTTATGAAGAGAAACGATATACAGATGTAAGAGAGATAATCATTTGT 3525
 Db 3651 AATTACAAACCGATTTATGAAGAGAAACGATATACAGATGTAAGAGAGATAATCATTTGT 3710
 Qy 3526 GAATATACAGAGGTTATGTAATTTACCACTACCACTACCACTGTTTATATGACAAAAGAA 3585
 Db 3711 GAATATACAGAGGTTATGTAATTTACCACTACCACTACCACTGTTTATGTAACAAAAGAA 3770
 Qy 3586 TTAGAATACCTTCCAGAAACCGATTAAGTTGAGATTGAGATTGAGAAACCGAAGGGAAG 3645
 Db 3771 TTAGAATACCTTCCAGAAACCGATTAAGTTGAGATTGAGATTGAGAAACCGAAGGGAAG 3830
 Qy 3646 TTTATTGTAGACGCGTGAATTTACTCTTTATCGAGGAATAG 3687
 Db 3831 TTTATTGTAGACGCGTGAATTTACTCTCTCATGGAAGAAATAG 3872

2031 ATTCAGGCGCTAGTGGAAATGGGAAGTGTATATAGATAAAATTTGAAATTTCCAGTT 2090
 1918 ACTGCAACCTTGAAGCAGAAATCGATTTAGAAAGCGCGAGAGCGGTGAATTCCTG 1977
 2091 ACTGCAACCTTGAAGCAGAAATCGATTTAGAAAGCGCGAGAGCGGTGAATTCCTG 2150
 1978 TTTTACTAATACCAATCCAAAGAGATTGAAACAGATGTGACAGATTATCATATTGATCAA 2037
 2151 TTTTACTAATACCAATCCAAAGAGATTGAAACAGATGTGACAGATTATCATATTGATCAA 2210
 2038 GTATCCAAATTTAGTGGCGTGTATCCGATGAAATTCCTGATGAAAGAGAGAAATTA 2097
 2211 GTATCCAAATTTAGTGGCGTGTATCCGATGAAATTCCTGATGAAAGAGAGAAATTA 2270
 2098 CTTGAGAAAGTGAATATGCGAAACGACTCAGTGTGATGAAAGAACTTTACTCCAGATCCA 2157
 2271 CTTGAGAAAGTGAATATGCGAAACGACTCAGTGTGATGAAAGAACTTTACTCCAGATCCA 2330
 2158 AACTTCATCCATCAATTAAGCAACCAAGACTTCAATCTACTAATGAGCAATCGAATTC 2217
 2331 AACTTCATCCATCAATTAAGCAACCAAGACTTCAATCTACTAATGAGCAATCGAATTC 2390
 2218 ACATCTATCCATGAACTCTGAACATGGATGGTGGGAAAGTGAGACATTTCAATCCAG 2277
 2391 ACATCTATCCATGAACTCTGAACATGGATGGTGGGAAAGTGAGACATTTCAATCCAG 2450
 2278 GAAGGAAATGACGTATTTAAAGGAAATTAACGTCACTACCGGGGACTTTTAATGAGTGT 2337
 2451 GAAGGAAATGACGTATTTAAAGGAAATTAACGTCACTACCGGGGACTTTTAATGAGTGT 2510
 2338 TATCCGAGCTTATTTATCAAAAATTAAGGAGAGTGGAAATTAAGCTTATATCTCGTAC 2397
 2511 TATCCGAGCTTATTTATCAAAAATTAAGGAGAGTGGAAATTAAGCTTATATCTCGTAC 2570
 2398 CAATTAAGAGGTATATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTATAAT 2457
 2571 CAATTAAGAGGTATATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTATAAT 2630
 2458 GCGAAAATCAATGAAATGATTTCCAGGTACCGAGTCCGATGGCGCTTTCCAGTTGAA 2517
 2631 GCGAAAATCAATGAAATGATTTCCAGGTACCGAGTCCGATGGCGCTTTCCAGTTGAA 2690
 2518 AGCCCAATCGAAGGTGGAGAACCGAATCGATCGCACCAATTTGATGAAATCGT 2577
 2691 AGCCCAATCGAAGGTGGAGAACCGAATCGATCGCACCAATTTGATGAAATCGT 2750
 2578 GATCTAGATTTCTTCGACAGATGGAGAAATGTCGGCATCATTTCCCATCATTTCTCT 2637
 2751 GATCTAGATTTCTTCGACAGATGGAGAAATGTCGGCATCATTTCCCATCATTTCTCT 2810
 2638 TTGGATATTCATATTTGATGCACAGACTTCATGAGAAATCTAGCGGTGGTGGTATTC 2697
 2811 TTGGATATTCATATTTGATGCACAGACTTCATGAGAAATCTAGCGGTGGTGGTATTC 2870
 2698 AAGATTAAGACGCGAGGAGTCTATGCAAGACTAGGGAATCTGGAATTTATTGAAGAGAAA 2757
 2871 AAGATTAAGACGCGAGGAGTCTATGCAAGACTAGGGAATCTGGAATTTATTGAAGAGAAA 2930
 2758 CAATTAATAGGAGACGACTCTCTGTTGAGAGAGACAGAGAAATTAAGAGAGACAAA 2817
 2931 CCATTAATAGGAGACGACTCTCTGTTGAGAGAGACAGAGAAATTAAGAGAGACAAA 2990
 2818 CGTGAATACTACAAATTTGAAACCAAAACGAGTATATACAGAGGCAAAAGAGCTGGAT 2877
 2991 CGTGAATACTACAAATTTGAAACCAAAACGAGTATATACAGAGGCAAAAGAGCTGGAT 3050
 2878 GCTTTATTTGATATCTCAATTAATAGATTTACAGCGGATCAAAACATTTGGCATGATT 2937
 3051 GCTTTATTTGATATCTCAATTAATAGATTTACAGCGGATCAAAACATTTGGCATGATT 3110
 2938 CATCGGCGAGATAAATTTGTTTCATCGAATTTGAGAGGCTTATCTGTGAGAAATTTCTGTT 2997

3111 CATCCGCGCAGATAAACTTTGTTTCATCGAATTCGAGAGCGGTATCTTTTTCAGAAATTTACCTGTT 3170
 2998 ATCCCGGCTGTAATGCGGAAAATTTTGAAGAAATTTAGAGGTGCATATATCATCTGCAATC 3057
 3171 ATCCAGGTGTAATGCGGAAAATTTTGAAGAAATTTAGAGGTGCATATATCATCTGCAATC 3230
 3058 TCCTTATACGATCGGAGAAATGTCGTTAAATAATGTTGATTTTAAATATGATTAATGATTAACATGT 3117
 3231 TCCTTATACGATCGGAGAAATGTCGTTAAATAATGTTGATTTTAAATATGATTAATGATTAACATGT 3290
 3118 TGAATGTAATAAGGGCATGTAGATGTACACAGAGCCATCACCGTTCCTGCTTGTATTC 3177
 3291 TGAATGTAATAAGGGCATGTAGATGTACACAGAGCCATCATCTGTTGACCTTGTATTC 3350
 3178 CCGAATGGAAGCAGAGTGTCAAGCAGTTCGCGTCTGTCCGGGCGTGGGTATATTC 3237
 3351 CCGAATGGAAGCAGAGTGTCAAGCAGTTCGCGTCTGTCCGGGCGTGGGTATATTC 3410
 3238 CTCGTTGTCACAGCGTACAAAGAGGGATATGGAAGGGCTGCGTAAACGATCCATGAAATC 3297
 3411 CTCGTTGTCACAGCGTACAAAGAGGGATATGGAAGGGCTGCGTAAACGATCCATGAAATC 3470
 3298 GAGAACAAATACAGACGAACTTAAATAATTTAAATACTGTGAAGAGAGGAAGTGTATCCCAACG 3357
 3471 GAGAACAAATACAGACGAACTTAAATAATTTAAATACTGTGAAGAGAGGAAGTGTATCCCAACG 3530
 3358 GATACAGGAACTGTATTAATGATTTACTGTCACCAAGGTACAGC-----AGTA 3405
 3531 GATACAGGAACTGTATTAATGATTTACTGTCACCAAGGTACAGCTGGATGCGCAGATGCA 3590
 3406 TGTAAATCCCGTAAATGCTGATATGAGGATGCAATGAAAGTTGATATACAGCATCTGTT 3465
 3591 TGTAAATCCCGTAAATGCTGATATGAGGATGCAATGAAAGTTGATATACAGCATCTGTT 3650
 3466 AATTACAAACCGACTTATGAAGAGAAACCTATACAGATGTACGAGAGATAATCATTTGT 3525
 3651 AATTACAAACCGACTTATGAAGAGAAACCTATACAGATGTGAAGAGATAATCATTTGT 3710
 3526 GAATATGACAGAGGGTATGTGAATTTATCCACCTACTACAGCTGGTTATATGACAAAAAGAA 3585
 3711 GAATATGACAGAGGGTATGTGAATTTATCCACCTACTACAGCTGGTTATATGACAAAAAGAA 3770
 3586 TTGAATATCTCCAGAACCGATTAAGTATGATGATGAGATTTGAGAGAAACGGAAGGGAAG 3645
 3771 TTGAATATCTCCAGAACCGATTAAGTATGATGATGAGATTTGAGAGAAACGGAAGGGAAG 3830
 3646 TTTTATTTGACAGCGTGGAAATTTACTCTTATCGAGGAATAG 3687
 3831 TTTTATTTGATAGTGGGAATTTACTCTTATCGAGGAATAG 3872

RESULT 12

AP363025

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

AP363025 4798 bp DNA linear BCT 07-MAY-2001
 Bacillus thuringiensis subsp. entomocidus delta-endotoxin CryIbA2
 (cryIbA2) gene, complete cds.
 AP363025
 AP363025.1 GI:13959050
 Bacillus thuringiensis serovar entomocidus
 Bacillus thuringiensis serovar entomocidus
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
 cereus group.
 1 (bases 1 to 4798)
 Mat Isa M.N., Abdullah, M.A.F. and Mahadi, N.M.
 Characterization of cryIb gene and its flanking regions cloned from
 Bacillus thuringiensis subsp. entomocidus HD-9
 Unpublished
 2 (bases 1 to 4798)
 Mat Isa M.N., Abdullah, M.A.F. and Mahadi, N.M.
 Direct Submission
 Submitted (21-MAR-2001) Center for Gene Analysis and Technology,
 Universiti Kebangsaan Malaysia, UKM Bangi, Selangor 43600, Malaysia

FEATURES

source Location/Qualifiers
 1..4798 /organism="Bacillus thuringiensis serovar entomocidus"
 /mol_type="genomic DNA"
 /strain="HD-9"
 /sub_species="entomocidus"
 /db_xref="taxon:1436"
 193..231
 promoter /note="BtII promoter"
 215..243 /note="BtI promoter"
 406..4107 /gene="cry1Ba2"
 406..415 /gene="cry1Ba2"
 /note="putative"
 421..4107
 /gene="cry1Ba2"
 /codon_start=1
 /transl_table=11
 /product="delta-endotoxin Cry1Ba2"
 /protein_id="AAK51084.1"
 /db_xref="GI:13959051"
 /translation="MISNRKNEHEINAVNSHSAQMDLLPDAIEDLSLCTAEGNNDIP
 FVSASTVQTGINAGRIILGVLGPFAGQASFSYFLVGLWPRGRDQWELFLEHVBQL
 INQITENARNTALRLQGLGDSFRAYQQSLEDLENRDDARTSVLHTQYIALBELDF
 LNAPLFAIRNQSVPLLMVYQAANLHLRLDASLFGSEFGLTSQBIQRYEYRQVER
 TRDYSYCVSWYNTGLNSLGTNAASWVRNQFRDLTLGLDLVALFPSYDRTYPI
 NTSAQLTREYVYDAIGATGVNMAWNNNAFSAIEAAAIRSPHLLDFLEQLIF
 SASRNSNTKMTYKRGHTLQSRPIGGGLNTSHGATNSINPVLRFPAERDVYRES
 YAGVLLGIYELIHGVTFRFNTPNQISDRGTANYSQPYSPGLQLKDSLELPP
 ETTFRPNYESYSHRLSHIGIILQSRVNVPIYVTHRSADRTNIGPNRIITQIPMKVAS
 ELPOGTTVNRGPGFTGLDILRNTGFGPIRVTVNGPLTQVRIGFVASTVDPDF
 VSRGTTVNRGPIRLTWNMSDELKYGNFVERAFTPTFTQIODIIRTSQGLSGNGE
 VYIDKIEIIPVATPEAYDLERAQAVNALFTNPNRLKTDVTDYHIDQVSNLVAC
 LSDFELCKRELLEKVKYAKRUSDERNLLQDPNFSINKQPDFISTNEOSNFTSIEH
 QSHGWGNSNITIQGNDVFNKYVTLPGTFNECYPTLYLQKIGSELKAYTRYQLR
 GYTESQDLIELYIRNAKHETLDVPCTESLWPLSVESPIRCGEBNRCAPHEWNP
 LDCSQRDEKCAHSHHSLDIDVGCTDLHENLGVVVFKIKTQEGHARLGNLEPDE
 KPLGLSALVRKAEKWRDREKLOLETKRVTVEAKEAVALFVDSOYDRLQADNI
 GMTAAADKLVRHIREAYLSLPLVPGVNAEIPFELSGHIIITALSIDARNVNGDPN
 NGLTGNVNRGVDVQSHHRSDLVPEWEAEVSQVRVCPGCGIILURVYAKSGYEGG
 CVTHIEINNTDLKPKRNEEVEYPTDGTNDYTAHQGTAGCAGACNRSNAGYEDA
 YEVDITTSVNRKPYEEETVYDVRDNHCEYDRGVNYPVPVAGYVTKLEYFPETDT
 VWIEIGTEGKFIQVDSVELLLMBE"
 repeat_region 4200..4238
 /rpt_type=inverted

ORIGIN

Query Match 85.0%; Score 3133.6; DB 1; Length 4798;
 Best Local Similarity 91.1%; Pred. No. 0;
 Matches 3373; Conservative 0; Mismatches 299; Indels 30; Gaps 3;
 1 TTGACTTCAATAGGAAATGAGATGAATTAATAATGCTTATCGATTCAGCTGTA 60
 421 TTGACTTCAATAGGAAATGAGATGAATTAATAATGCTTATCGATTCAGCTGTA 465
 61 TCGAATCATTTCCACAAATGATCTATCACAGATGCTCGTATTAGGATTTCTTTGT 120
 466 TCGAATCATTTCCGACAAATGATCTATTAACAGATGCTCGTATTAGGATTTCTTTGT 525
 121 ATAGCCGAGGGGAATAATCAATCCATTGTTAGCGCATCAACAGTCCAAACGGGTATT 180
 526 ATAGCCGAGGGGAATAATGATCCATTGTTAGCGCATCAACAGTCCAAACGGGTATT 585
 181 AACATAGCTGGTAGAATCTAGGTGATATTAGCGGTACCGTTTGTGACAAATAGCTAGT 240
 586 AACATAGCTGGTAGAATCTAGCGGTATTGCGCGTACCGTTTGTGACAAATAGCTAGT 645
 241 TTTTATAGTTTTCTTTGTTGTAATTATGCCCCCGCGGACAGATCAGTGGGAAATTTTC 300
 646 TTTTATAGTTTTCTTTGTTGTAATTATGCCCCCGCGGACAGATCAGTGGGAAATTTTC 705

301 CTAGAACATGTGCAACAACTTATAAATCAACAAATTAACAGAAATGCTAGAAATACGGCA 360
 706 CTAGAACATGTGCAACAACTTATAAATCAACAAATTAACAGAAATGCTAGAAATACGGCT 765
 361 CTTGCTCGATTACAAGGTTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACCTTGAGAT 420
 766 CTTGCTCGATTACAAGGTTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACCTTGAGAT 825
 421 TGGCTAGAAACCGTGATGATGCAAGAACGAGAAGTGTCTTTATATACCCCAATATATAGCC 480
 826 TGGCTAGAAACCGTGATGATGCAAGAACGAGAAGTGTCTTTATATACCCCAATATATAGCT 885
 481 TTAGAACTTGATTTTCTTAATGCGATGCGCTTTTTCGAATTAAGAAACCAAGAGTTTCCA 540
 886 TTAGAACTTGATTTTCTTAATGCGATGCGCTTTTTCGAATTAAGAAACCAAGAGTTTCCA 945
 541 TTATTAATGATGATGCTCAAGCTGCAAAATTTACACTATTTATTTAGAGATGCTCTCT 600
 946 TTATTAATGATGATGCTCAAGCTGCAAAATTTACACTATTTATTTAGAGATGCTCTCT 1005
 601 CTTTGTGAGTGAATTTTGGGCTTTACATCGCAGGAAATTCACGTTTATTTAGAGCGCAA 660
 1006 CTTTGTGAGTGAATTTTGGGCTTTACATCGCAGGAAATTCACGTTTATTTAGAGCGCAA 1065
 661 GTGGAACAAACGAGAGATTTATTCGACTATTCGCTAGATGGTATTAATACAGCTTAAAT 720
 1066 GTGGAACAAACGAGAGATTTATTCGACTATTCGCTAGATGGTATTAATACAGCTTAAAT 1125
 721 AGCTTGAGAGGGAACAATCGCGAAGTGGTGGTATTAATCAATTCCTAGAGATCTA 780
 1126 AGCTTGAGAGGGAACAATCGCGAAGTGGTGGTATTAATCAATTCCTAGAGATCTA 1185
 781 ACCTTGAGGGAATTTAGATCTAGTGGCACTATTCGCTAGATGGTATTAATCAATTCCTAGAGATCTA 840
 1186 ACCTTGAGGGAATTTAGATCTAGTGGCACTATTCGCTAGATGGTATTAATCAATTCCTAGAGATCTA 1245
 841 ATAAATACGAGTCTCAGTTTAAACAGGGAAGTTTATACAGCGCAATTTGAGCAACAGGG 900
 1246 ATAAATACGAGTCTCAGTTTAAACAGGGAAGTTTATACAGCGCAATTTGAGCAACAGGG 1305
 901 GTAAATATGCAAGTATGATGCTGATTAATAATGCACTTCCTGTTTCCTGCTATAGAG 960
 1306 GTAAATATGCAAGTATGATGCTGATTAATAATGCACTTCCTGTTTCCTGCTATAGAG 1365
 961 ACTGCGGTTTATCGAAGCCGCACTCTACTTGTATTTCTAGAACAACTTACAAATTTTATG 1020
 1366 GCTGCGGCTATCGAAGCCGCACTCTACTTGTATTTCTAGAACAACTTACAAATTTTATG 1425
 1021 ACTTCATCAGATGGAGTGTCTAGGCATATGACTTACTGGCGGGGGGCAACAATTCAA 1080
 1426 GCTTCATCAGATGGAGTGTCTAGGCATATGACTTACTGGCGGGGGGCAACAATTCAA 1485
 1081 TCTCGGCAATAGAGCGGATTAATAATCCTCAACGATGGGTCTACCAATCTCTTATT 1140
 1486 TCTCGGCAATAGAGCGGATTAATAATCCTCAACGATGGGTCTACCAATCTCTTATT 1545
 1141 AATCCTGTAGATTTATCTTCTCTCTCGAGAGCTATATTGGAATCATATGACAGGA 1200
 1546 AATCCTGTAGATTTATCTTCTCTCTCGAGAGCTATATTGGAATCATATGACAGGA 1605
 1201 GTGCTTCTATGGGAAATTTACCTTGAACCTATTTCATGGGTGCTCTACTGTTAGATTTAAT 1260
 1606 GTGCTTCTATGGGAAATTTACCTTGAACCTATTTCATGGGTGCTCTACTGTTAGATTTAAT 1665
 1261 TTTTAGGAACCTCAGAAATCTTTTGAAGAGGTACTGCTAACTATAGTCAACCTATGAG 1320
 1666 TTTTAGGAACCTCAGAAATCTTTTGAAGAGGTACTGCTAACTATAGTCAACCTATGAG 1725
 1321 TCACCTGGGCTTCAATTTAAAGATTTCAGAAATTCGAATTTACCAACAGAAACCAAGAGAG 1380
 1726 TCACCTGGGCTTCAATTTAAAGATTTCAGAAATTCGAATTTACCAACAGAAACCAAGAGAG 1785
 1381 CCAATTTAGAAATCATATAGTCTAGGTTATCTCAGATAGGCTCATTTTCAATCTAGG 1440

[illegible]

D5	D5	2866	GC	GAACATGAAACATTTGGATGTTCCAGGTACCGAGTCCCTATGCGCGCTTTCAGTTGAA	2922
Q5	Q5	2518	AG	CCCAATCGGAAGTGC GGAGAACCGAAATCGATGCGCACCAATTTTGAATGGAATCCT	2577
D5	D5	2926	AG	CCCAATCGGAAGTGC GGAGAACCGAAATCGATGCGCACCAATTTTGAATGGAATCCT	2985
Q5	Q5	2578	GA	CTAGATTTGTTCTCTCAGAGATGGAGAAAATGTGCGCATCATTTCCCATCATTTCTCT	2637
D5	D5	2986	GA	CTAGATTTGTTCTCTCAGAGATGGAGAAAATGTGCGCATCATTTCCCATCATTTCTCT	3045
Q5	Q5	2638	TT	GGATATTCGATATTGGATGCACAGACTTCGATCAGAAATCTAGCGGTGTGGGTGGTATTC	2697
D5	D5	3046	TT	GGATATTCGATATTGGATGCACAGACTTCGATCAGAAATCTAGCGGTGTGGGTGGTATTC	3105
Q5	Q5	2698	AA	GATTAAAGACGCAAGGAAGTCAATGCAAGACTAGGGAAATCGAAATTTATTGAAGAGAAA	2757
D5	D5	3106	AA	GATTAAAGACGCAAGGAAGTCAATGCAAGACTAGGGAAATCTGGAAATTTATTGAAGAGAAA	3165
Q5	Q5	2758	CC	ATTATTAGGAGNAGCACTGTCTCGTGTGAAGAGCGAGAGAAAAAATGGAGAGACAAA	2817
D5	D5	3166	CC	ATTATTAGGAGNAGCACTGTCTCGTGTGAAGAGCGCGAGAGAAAAAATGGAGAGACAAA	3225
Q5	Q5	2818	CG	TGAAAAACTCAATTTGGAAAAACAACAGAGTATATACAGAGGCAAAAGAGCTGTGGAT	2877
D5	D5	3226	CG	TGAAAAACTCAATTTGGAAAAACAACAGAGTATATACAGAGGCAAAAGAGCTGTGGAT	3285
Q5	Q5	2878	GC	TTTATTTGTAGATTTCTCAATATAATAGATTACAAGCGGATACAAACATTTGCATGATT	2937
D5	D5	3286	GC	TTTATTTGTAGATTTCTCAATATGATAGATTACAAGCGGATACAAACATTTGCATGATT	3345
Q5	Q5	2938	CA	TGCGCGAGATAAACTTTGTTTCATTCGAATTCGAGAGGCTTATCTGTACAGAAATTTCTGTT	2997
D5	D5	3346	CA	TGCGCGAGATAAACTTTGTTTCATTCGAATTCGAGAGGCGTATCTTTACAGAAATTTACCTGTT	3405
Q5	Q5	2998	AT	CCGGGTCTTAATCGGAAATTTTGAAGAAATTAGAAGTTCGCACTTATCACTGCAATC	3057
D5	D5	3406	AT	CCAGGTGTAAATCGGAAATTTTGAAGAAATTAGAAGGTACACTTATCACTGCAATC	3465
Q5	Q5	3058	TC	CCCTATACGATCGAGAAATGTCTGTAAAAATGGTGATTTTAAATAGGATTTAGCATGC	3117
D5	D5	3466	TC	CTTATACGATCGAGAAATGTCTGTAAAAATGGTGATTTTAAATAGGATTTAACTGT	3525
Q5	Q5	3118	TG	GAATGTAAAAGGCAATGTAGATGTATACAAACAGAGCCATCACCGTTCTGTCTTGTATC	3177
D5	D5	3526	TG	GAATGTAAAAGGCAATGTAGATGTATACAAACAGAGCCATCATCTGTCTGACCTTGTATC	3585
Q5	Q5	3178	CC	AGAATGGGAACGACAAGGTGTACAAAGCAAGTTCCGGTCTGTCCGGGGCGTGCTATATC	3237
D5	D5	3586	CC	AGAATGGGAACGACAAGGTGTACAAAGCAAGTTCCGGTCTGTCCGGGGGTGCTCTATC	3645
Q5	Q5	3238	CT	CCGTCTCACAGCGTACAAAGGGGATATGGAGAGGTTGTGTAAACGATCCATGAATC	3297
D5	D5	3646	CT	CCGTCTCACAGCGTACAAAGAGGGATATGGAGAGGCTGCGTTAACGATCCATGAATC	3705
Q5	Q5	3298	GAG	AACAATACAGACGAACCTAAATTTTAAAACTGTGTGAAGAGAGGAAGTGTATCCAACG	3357
D5	D5	3706	GAGA	ACAATACAGACGAACCTAAATTTTAAAACTGTGTGAAGAGAGGAAGTGTATCCAACG	3765
Q5	Q5	3358	GAT	CACAGACGCTGTAAATGATTTACTTCGACACCAAGGTACAGC-----AGTA	3405
D5	D5	3766	GAT	CACAGACGCTGTAAATGATTTACTTCGACACCAAGGTACAGCTGGATCGCAGATGCA	3825
Q5	Q5	3406	TGT	AAATTTCCCGTAAATGCTCGGATATGAGGATGCATATGAAGTTGATACTACAGCATCTGTT	3465
D5	D5	3826	TGT	AAATTTCCCGTAAATGCTCGGATATGAGGATGCATATGAAGTTGATACTACAGCATCTGTT	3885
Q5	Q5	3466	AA	TACAAACCGACTTATGAAAGAGAAACGTTATACAGATGTACGAAGATATCATTTGT	3525
D5	D5	3886	AA	TACAAACCGACTTATGAAAGAGAAACGTTATACAGATGTACGAAGATATCATTTGT	3945
Q5	Q5	3526	GA	TATGACAGAGGGTATGTAATTTATCCACACTACAGCTGCTGTTTATATGACAAAGAA	3585
D5	D5	3946	GA	TATGACAGAGGGTATGTCATTTATCCACAGTACCACTGCTGGTGTATGTGACAAAGAA	4005

Qy	3586	TTAGAATATCTTCCAGAGAAACCGATATGATTGAGATTGAGAGAAACGAGAGGGAAG	3645
Db	4006	TTAGAATATCTTCCAGAGAAACAGATACAGTATGATTGAGATTGAGAGAAACGAGAGGGAAG	4065
Qy	3646	TTTATTGTAGACAGCGTGGAAATTACTCCTTATTCGAGGAAATAG	3687
Db	4066	TTTATTGTAGATAGCGTGGAAATTACTCCTCATTCGAGAAATAG	4107

RESULT 13	
LOCUS	BACCRYIE
DEFINITION	Bacillus thuringiensis crystal protein (cryI ET5) gene, complete cds.
ACCESSION	L32020
VERSION	L32020.1
KEYWORDS	GI:474893 crystal protein.
SOURCE	Bacillus thuringiensis
ORGANISM	Bacillus thuringiensis
REFERENCE	Bacteria, Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
AUTHORS	1 (bases 1 to 3934) Donovan,W.P.
JOURNAL	Unpublished (1994)
COMMENT	Original source text: Bacillus thuringiensis (strain EG 5847) DNA.
FEATURES	Location/Qualifiers
SOURCE	1..3934

```

gene
cds
/db_xref="taxon:11428"
67..3756
/gene="cryI ETS"
67..3756
/gene="cryI ETS"
/function="insect toxin"
/codon_start=1
/transl_table=11
/product="crystal protein"
/protein_id="AAA22344.1"
/db_xref="GI:474894"
/translation="MTSNRNKEEINIALSIPTVSNPSTQMNIISPARIEDISLCVAEFL
NNIDPFVASTVTGNIAGRIILVGLVFPAGLQAFYFVLGELMPSGRDWEIFLE
HVEDLIRQVOTENTRNTARLEBGLGRYSYQOALETMDLRNDRARSIIILERYVA
LEDITQATPLFRINEEYPLMVYAAQANLHLLLRDASLFGSEWMASSIDVNOYQYD
EQIYRTEESYHNCHVQMYNTGLNLRGTNAESWLRYNQFRDRLTLGVLDLVPRLPILDF
RVPYINSAQLTRPIYTDPIGRNAPSGPASTWFNNAFSPSAIEAALFPRPHLIDF
PEQLTIYSASSRWSSTQHNWYVCHRLNFRPIQGLTNTSTQGLTNTSINPTVLQFTS
RDVYTESNAGTNLITPVPNGVWPAFNFINFQPIYERGAITYSQYGVGQLQFDS
ETELPEPTTERPNYESYHRLLSHLGLIIGNTLRAPVYSWTHRSADRNTIIGNRNIQI
PLVKAHLNHSVTGVTGGPGTGGIILRTNLTGTLIRNLINVLPSQRVYRILRYAST
TDLQFTRINGTIVTGNFSRTWMRGDNLREYSTRFAGSTPFNFNAQSTFTFLGAQS
NMQVEYIDRVEVPEAVLPFEASVYDLERAQKAVNALPTSTNPRLLKTDVTDYHDQVS
FNVACLSDRELPEKVLPEKAVKRLSDERNLDPNFTIYSGLQSPASIDQGNF
PSNLSEHGWSGANVYLTQEGNDVFENAVTLPTFNCPYNYLVKQIGSESLKAVT
RYQLRLEYSQDLQLEIYLIRYNAKHETLVBGTDSLWPLSVESPIGRCEPNRCADHF
EWNFDQSCDRGECARHSHFTDLIDVGCTDLHNLGWWVFKIKIQEYARLNL
FTLEEKPLGEALSRVKRAEKKWRDKREKLKTVYTEAKEAVDALFVDSYDQQLQ
EPTNLGMILHAADKLVRHIREAYLSELPVPGNAEYFEELEGHIITAMSLDARNVVK
NGDFNGLGTCWNVYKHDVQOHSRSDLVIPWEAEVSAQVRCVPCQGYILRVTAYKE
GYEGCVGTHEIEINNTDELKFNCEBEEVYPTDGTCDNYTAHQGTAAACNRSNAGED
AYEVDVTTAEYNYKPYEEETVTDVRDNDHCEYDRGVVYPPVPAGVYVYKLEYPETD
TWIIEIGETEGKFI VDSVLLMEE"

```

ORIGIN

Query Match	72.9%;	Score	2687.2;	DB	1;	Length	3934;
Best Local Similarity	83.5%;	Pred. No.	0;				
Matches	308;	Conservative	0;	Mismatches	593;	Indels	15;
Gaps	3;						

QY	1	TTGACTTCAAATAGGAAAAATGAGATGAATATATAAATCGTTTATCGATTCCAGCTGTA	60
Db	67	TTGACTTCAAATAGGAAAAATGAGATGAATATATAAATCGTTTATCGATTCCACCGTA	126

Qy	61	TCGAATCATTTCCACAAATATTCATACACAGATGCTCGTATTTGAGGATTTCTTTGTGT	120
Db	127	TCGAATCCCTTCCACGAAATGAATCTATCACAGATGCTCGTATTTGAAGATATGCTTGTGT	186
Qy	121	ATAGCCGAGGGGAATATATCAATCCACTTGTGTAGCGCATCAACAGTCCAAACGGGTATT	180
Db	187	GTAGCCGAGGTGAACAATATTTGATTCATTTGTGTAGCGCATCAACAGTCCAAACGGGTATA	246
Qy	181	AACATAGCTGGTAGAATATCTAGGTGTATTAGGCGTACCGTTTGTCTGACAAAATAGCTTAGT	240
Db	247	AACATAGCTGGTAGAATATTTGGCGTATTAGGTGTGCGGTTTGTCTGACAACTAGCTTAGT	306
Qy	241	TTTTATAGTTTTCTTGTGTGAATATATGCCCCCGCGCAGAGATCAGTGGGAAATTTTC	300
Db	307	TTTTATAGTTTTCTTGTGGGAAATATATGCCCTAGTGGCAGAGATCCATGGGAAATTTTC	366
Qy	301	CTAGAACATGTCGAACAACTTTATTAATCAACAATAACAGAAAATGCTAGGAATACGGCA	360
Db	367	CTGGAACATGTGAACAACTTTATTAAGACAACTAGATTAACAGAAAATCTAGGAATACGGCT	426
Qy	361	CTTGTCTGATTACAAGTTTTAGGAGATTCTTWTAGAGCCTATCAACAGTCACTTTGAAGAT	420
Db	427	ATTGTCTGATTAGAGTCTAGGAGAGGCTATAGATCTTTACAGCAGGCTCTTGAACCT	486
Qy	421	TGGCTAGAAAACCGTGATGATGCAAGAACGAGAAGTTCTTTTATACCCAAATATATAGCC	480
Db	487	TGGTTAGATAACCGAAATGATGCAAGATCAAGAAGCATTTCTTTGAGCGCTATTTGTCT	546
Qy	481	TTAGAACTTGATTTTTCTTAATGCGATGCGCTTTTCGCAATTAGAAAACCAAGAGTTCCA	540
Db	547	TTAGAACTTGACATTTACTCTGCTATACCGCTTTTTCAGAAATACGAAATGAAGAAGTTCCA	606
Qy	541	TTATTAATGTTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGCTCT	600
Db	607	TTATTAATGTTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGAGCTATCC	666
Qy	601	CTTTTTGGTAGTGAATTTGGGCTTTACATCGCAGGAAATTCACGTTATTATGAGCGCCAA	660
Db	667	CTTTTTGGTAGTGAATGGGGATGGCATCTTCCGATGTTAAACCAATATTACCAAGAACAA	726
Qy	661	GTGGAAACAACGAGAGATTATTCGGACTATTTCGGTAGAATGGTATTAATACAGTCTTAAT	720
Db	727	ATCAGATATACAGAGAAATATTTCAACCAATGGGTCAATGGTATTAATACAGGGCTAAAT	786
Qy	721	AGCTTCGAGGGGCAAAATCGCGAAGTGTTGGTGGCTTTAATACTCAATTCCTGAGAGACTA	780
Db	787	AACTTAAGGGGCAAAATGCTGAAGTTGGTGGGTATTAATCAATTCCTGAGAGACTTA	846
Qy	781	ACGTTAGGGGTAATAGATCTAGTGGCACTATTTCCCAAGCTATGACATCGCACTATFCCA	840
Db	847	ACGTTAGGGGTAATAGATTTAGTAGCCCTATTTCCCAAGCTATGATCTCGCACTTATFCCA	906
Qy	841	ATAAATACGAGTGCTCAGTTAAACAAGGGAAGTTTATACAGACGGAATTTGAGCAACAGGG	900
Db	907	ATCAATACGAGTGCTCAGTTAAACAAGAGAAATTTATACAGATCCAAATTTGGGAGAACAAAT	966
Qy	901	GTAAAT-----ATGCCAGTATGAATTGGTATAATAATATGCACCTTCGTTTTCCGCT	954
Db	967	GCACCTTCAGGATTTGCAAGTAGAAATTTGGTTTAAATAANTAGCACCATCGTTTTCTGCC	1026
Qy	955	ATAGAGACTCGGGTTATTCGAAGCCCGCACTACTTGAATTTCTTAGAAACAATTTACAATT	1014
Db	1027	ATAGAGGCTGCCATTTTCAGGCCCTCGCATCTACTTGAATTTTCAGAAACAATTTACAATT	1086
Qy	1015	TTTAGCATTTTCATCAAGATGGATGCTATAGGCAATATGACTTATCGCGGGGGCACACA	1074
Db	1087	TACAGTGCATCAAGCCGTTGGAGTAGCACTCAACATATGAATTTATTTGGGTGGGACATAGG	1146
Qy	1075	ATTCAATCTCGGCCAATAGGAGCGGATTAATAATACCTCAACGGATGGGTCTACCA---AT	1131
Db	1147	CTTAATCTTCGCCCAATAGGAGGGAATTAATAATCTCAACAAAGACTCTAATAAT	1206

1132 ACTTCTATTAACTCTGTGAAGATTATCAATCTTCTCTCGAGACGTATATTGGACGTGAATCA 1191
1207 ACTTCAATTAATCTGTAAACATTTACAGTTTACGTTCTCGAGACGTTTATAGAACAGATCA 1266
1192 TATGACGAGTCTCTATGGGGAATTTACTTGAACCTATTCTATGATGTCCTTACTGTT 1251
1267 AATGACGGACAAATAT-----ACTAATTTACTACTCTGTGATGGAGTACCTTGGGCT 1320
1252 AGATTTAATTTTATAGAAACCTTCAGAAATCTTTTGAAGAGGTACTGCTAACTATAGTCAA 1311
1321 AGATTTAATTTTATAAACCTTCAGAAATTTATGAAAGAGCGCCACTACCTACAGTCAA 1380
1312 CCCTATGAGTCACTTGGGCTTCAATTTAAAGATTTCAGAAACTGAAATTTACCAACGAAACA 1371
1381 CCGTATCAGGAGTGGGATTTCAATTTATTTGATTTGATTTTCAAGAACTGAATTTACCACAGAAACA 1440
1372 ACAGAACGCAAAATTAATGAATCATATAGTCATAGTTTATCTCACATAGGCTCATTTCA 1431
1441 ACAGAACGCAAAATTAATGAATCATATAGTCATAGTTTATCTCATATAGGACTAATCATTA 1500
1432 CAATCTAGGCTGCATGTACCAGTATATTTCTTGGACGCAACGCTAGTCAGATCGTACAAAT 1491
1501 GGAACACATTTGAGAGCACCACTATCTTGGACGCACTCGTAGTCAGATCGTACGAAT 1560
1492 ACCATTAGTTTCAGATAGCATAAACAAATACCAATTTGGTAAATCATTTCAACCTTAATCA 1551
1561 ACGAATGGACCAATAGAAATTAACAAATACCAATTTGGTAAAGCACTGAATCTTCATTCA 1620
1552 GGTACCTCTGTAGTCAGTGGCCAGGATTTACAGGAGGGATTAATCCGAACTAACGTT 1611
1621 GGTGTTACTGTGTGGAGGCGCAGGATTTACAGTGGGATATCTTCTCGTAGAACAAAT 1680
1612 AATGGTAGTGTACTAAGTATGGGTCTTAAATTTTAAATTAATACATTTACAGCGGTATCGC 1671
1681 ACGGTACATTTGGAGATATACGATTAATTAATTTGTGCGCATTTATCCCAAGATATCGC 1740
1672 GTGAGAGTTCGTTATGCTGTTCTCAACAAATGCTCTGAGGGTAACTGTGCGAGGAGT 1731
1741 GTAAGGATTCGTTATGCTTCTACTACAGATTTTACAATTTTTCACGAAATTAATGGAACC 1800
1732 ACTACTTTTGCATCAAGATTTCCCTAGTACTATGATGCAATGATGCTTTGACATCTCAA 1791
1801 ACTGTTAATATTTGTTAATTTCTCAAGAACTATGAATAGGGGGATTAATTTAGAAATATA 1860
1792 TCATTTAGATTTGCAGAAATTTCTGTAGGTATTTAGTGCACTGCGCAGTCAAACTGCTGGA 1851
1861 AGTTTTAGAACTGCAGATTTTAGTACTCTCTTTTAAATTTTAAATGCCCAAGCACATTC 1920
1852 ATAAGTATAAGTAAATATGACAGGTAGACAAACGTTTTCACTTTGAATAAATTTGAATTCAT 1911
1921 ACATGGGTCTCAGAGTTTTTCAATCAGGAAGTTTATATAGATAGAGTCGAATTTGTT 1980
1912 CCAATTTACTGCAACCTTCGAAGCAGAAATACGATTTAGAAAGGGCGCAAGAGGCGGTGAAT 1971
1981 CCAGCAGAGTAACTTTGAGGCGAGATATGATTTAGAAAGAGCACAAAGGGCGGTGAAT 2040
1972 GCTCTGTTTACTAATACGAATCCAGAGATTTGAAACAGATGTGACAGATTTATCATATT 2031
2041 GCTCTGTTTACTTCTACAAATCCAGAGATTTGAAACAGATGTGACAGATTTATCATATT 2100
2032 GATCAAGTATCCAAATTTAGTGGGCTGTTTATCGGATGAATTTCTGCTTAGATGAAAGAGA 2091
2101 GACCAAGTGTCCAATATGTTGGCATGTTTATCAGATGAATTTTGTGGATGAGAGCGA 2160
2092 GAATTTCTGAGAAAGTGAATATGCGAAACGACTCAGTGAATGAAGAACTTACTCCAA 2151
2161 GAATTTATTTGAGAAAGTGAATATGCGAAGCACTCAGTGAATGAAGAACTTACTCCAA 2220
2152 GATCCAACTTCATCATCAATCAATAGCAACCGACTTTCATCTATCTAATAGGCAATCG 2211
2221 GATCCAACTTCATCATCTCAGTGGCAATTAAGTTTCGATCCCATCGATGACATCA 2280
2212 AATTTACATCTATCCATGAACAACTCTGAACATGGATGGTGGGAGTGGAGAACATTTACA 2271

2281 AACTTCCCTCTATTAATAGCTATCTGAACATGGATGGTGGGGAAGTGCGAATTTACC 2340
2272 ATCCAGGAAGGAATGACGTATTTAAAGAAATTCAGTCACTACCTACCGGAGCTTTTAAT 2331
2341 ATTCAGGAAGGAATGACGTATTTAAAGAAATTCAGTCACTACCGGAGCTTTTAAT 2400
2332 GAGTGTATCCGACGTATTTATATCAAAAAATAGGAGAGTCGGAATTTAAAGCTTATACT 2391
2401 GAGTGTATCCAAATTTATATCAAAAAATAGGAGAGTCAGAAATTTAAAGCTTATACG 2460
2392 CGCTACCAATTAAGAGGATATTTGAAGTAGTCAAGATTTAGAGATATATTTGATTCGT 2451
2461 CGCTATCAATTAAGAGGATATTTGAAGTAGTCAAGATTTAGAGATATTTTAATTCGT 2520
2452 TATAATGCGAAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTCA 2511
2521 TACAAATGCAAGCATGAACATTTGGATGTTCCAGGTACCGATTTCCCTATGCGCGCTTTCA 2580
2512 GTTGAAGGCCAATCGGAAGGTGCGAGAACCGAATCGATGCGCACCACTTTTGAATGG 2571
2581 GTTGAAGGCCAATCGGAAGGTGCGAGAACCGAATCGATGCGCACCACTTTTGAATGG 2640
2572 AATCCTGATCTAGATTTCTCGAGAGATGGAAGAAATTTGGGATCATTTCCCATCAT 2631
2641 AATCCTGATCTAGATTTCTCGAGAGATGGAAGAAATTTGGGATCATTTCCCATCAT 2700
2632 TTCTCTTTGGATATTCATATTTGGATGCGACAGCTTTGATCAGAAATCTGAGCGTGTGGTG 2691
2701 TTCACTTTGGATATTCATATTTGGGTGCAACAGCTTTGATGAGAACTTGGCGTGTGGTG 2760
2692 GTATTCAAGATTAAAGACGCGAGGAGTCTATGCAAGACTAGGGAATCTGGAATTTATTTGAA 2751
2761 GTATTCAAGATTAAAGACGCGAGGAGTCTATGCAAGACTAGGGAATCTGGAATTTATTTGAA 2820
2752 GAGAAACCTATTATGAGAGAGCACTGTCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 2811
2821 GAGAAACCTATTATTTGAGAGAGCACTGTCTGCTGTGAGAGAGAGAGAGAGAGAGAGAG 2880
2812 GACAAACGTTGAAACCTACAAATTTGAAACCAAAACGAGTATATACAGAGGCAAAAGAGCT 2871
2881 GACAAACGTTGAAACCTACAAATTTGAAACCAAAACGAGTATATACAGAGGCAAAAGAGCT 2940
2872 GTGGATGCTTTATTTGTAGATTTCTCAATATAATAGATTTACAAGCGGATACAAACATTTGGC 2931
2941 GTGGATGCTTTATTTGTAGATTTCTCAATATGATCAATTTACAAGCGGATACAAACATTTGGC 3000
2932 ATGATTCATGCGCAGATAACTTTGTTTCATCGAATTTGAGAGGCTTATCTGTGAGAAATTA 2991
3001 ATGATTCATGCGCAGATAACTTTGTTTCATCGAATTTGAGAGGCTTATCTTTTCAAGAAATTA 3060
2992 TCTGTTATCCCGGTGTAAATGCGGAATTTTGAAGATTTAGAGGTCGCAATTTACCT 3051
3061 CCGTGTATCCCGGTGTAAATGCGGAATTTTGAAGATTTAGAGGTCGCAATTTACCT 3120
3052 GCAATCTCCCTATACCATGCGAGAAATGTCGTTTAAATAATGCTGATTTTAATAATTTGATTA 3111
3121 GCAATGCTCTTATACCATGCGGAATTTGCTTTAAATAATGCTGATTTTAATAATTTGATTA 3180
3112 GCAATGCTGGAATTTAAAGGCAATGTAGATGTACAACAGAGCCATCACCTTTCTGTCTCT 3171
3181 ACATGTTGGAATGTAAAGGCAATGTAGATGTACAACAGAGCCATCATCTTTCTGACCTT 3240
3172 GTTATCCAGATGGAAGGCAATGTAGATGTACAACAGAGGTCGCTGCTGCGGCGCGTGGC 3231
3241 GTTATCCAGATGGAAGGCAATGTAGATGTACAACAGAGGTCGCTGCTGCGGCGCGTGGC 3300
3232 TATATCTCTGCTCACAGGTACAAAGAGGATATGAGAGGCTGTGTGTAACGATCCAT 3291
3301 TATATCTCTGCTCACAGGTACAAAGAGGATATGAGAGGCTGTGTGTAACGATCCAT 3360
3292 GAAATCGAGAACAAATACAGACGAATCTTAAATAATTTGAGAGAGGAGGAGTGTAT 3351


```
Db 3361 GAAATCGAGAACAAATACAGACGAACTAAAATTTAAAAAATCTGAGAAAGAGGAAGTGTAT 3420
Qy 3352 CCACGGATACAGAAAGTGTATGATTTATCTACATGCGACACCAAGGTACAGCAGTATGTAAT 3411
Db 3421 CCACGGATACAGAAAGTGTATGATTTATCTACATGCGACACCAAGGTACAGCAGTATGTAAT 3480
Qy 3412 TCCCGTAATGCTGGATATGAGGATGCAATATCAAGTTGATATCAAGTCTACAGCATCTGTTAATTAC 3471
Db 3481 TCCCGTAATGCTGGATATGAGGATGCAATATCAAGTTGATATCAAGTCTACAGCATCTGTTAATTAC 3540
Qy 3472 AACCCGATTTATGAAGAAAGAAAGTATACAGATGACGAAGAGATATCAATGTAATAT 3531
Db 3541 AAACCCGATTTATGAAGAAAGAAAGTATACAGATGACGAAGAGATATCAATGTAATAT 3600
Qy 3532 GACAGAGGATGTAATATCCACCACTACCAAGTTCAGATTGAGAAACCGAAGGAAAGTTTATT 3591
Db 3601 GACAGAGGATGTAATATCCACCACTACCAAGTTCAGATTGAGAAACCGAAGGAAAGTTTATT 3660
Qy 3592 TACTTCCAGAAACCGATAAGTATGATTCAGATTGAGAAACCGAAGGAAAGTTTATT 3651
Db 3661 TACTTCCAGAAACAGATAAGTATGATTCAGATTGAGAAACCGAAGGAAAGTTTATT 3720
Qy 3652 GTAGACAGCTGGAATTTACTCTCTATGAGAGATAG 3687
Db 3721 GTAGATAGCTGGAATTTACTCTCTATGAGAGATAG 3756

RESULT 14
LOCUS AR068828 3934 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 3 from patent US 5854053.
ACCESSION AR068828
VERSION AR068828.1 GI:6001035
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3934)
AUTHORS Donovan, W. P. and Gonzalez, J. M., Jr.
TITLE Bacillus thuringiensis bacteria
JOURNAL Patent: US 5854053-A 3 29-DEC-1998;
FEATURES Location/Qualifiers
source 1..3934
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 72.9%; Score 2687.2; DB 6; Length 3934;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 3086; Conservative 0; Mismatches 593; Indels 15; Gaps 3;
Qy 1 TTGACCTTCAAAATAGGAAATGGAATGAAATATTAATGCTTTATCGATTCCAGCTGTA 60
Db 67 TTGACCTTCAAAATAGGAAATGGAATGAAATATTAATGCTTTATCGATTCCAAACGGTA 126
Qy 61 TCGAATCATTTCCACACAAATGGATCTATACCCAGATGCTCGTATGAGGATTTCTTTGT 120
Db 127 TCGAATCTTTCCAGCAAAATGAATCTATACCCAGATGCTCGTATGAGGATTTCTTTGT 186
Qy 121 ATAGCCGAGGGGAATATATCAATCCACTTTGTTAGCCGATCAACAGTCCAAAACGGGATT 180
Db 187 GTAGCCGAGGGGAATATATCAATCCACTTTGTTAGCCGATCAACAGTCCAAAACGGGATA 246
Qy 181 AACATAGCTGTAGAAATGATAGGTGATATAGGCGTACCGTTTCTGTCGACAAATAGCTAGT 240
Db 247 AACATAGCTGTAGAAATGATAGGTGATATAGGCGTACCGTTTCTGTCGACAAATAGCTAGT 306
Qy 241 TTTTATAGTTTCTTTGTTGTTGTAATATGCGCCCGCGCAGAGATCAGTGGGAAATTTTC 300
Db 307 TTTTATAGTTTCTTTGTTGTTGTAATATGCGCCCGCGCAGAGATCAGTGGGAAATTTTC 366
Qy 301 CTAGAACATGTGAAACAACTTATTAATCAACAAATAACAGAAATGCTAGGAATACGGCA 360
Db 301 CTAGAACATGTGAAACAACTTATTAATCAACAAATAACAGAAATGCTAGGAATACGGCA 360
```

```
Db 367 CTGGAACATGTGAAACAACTTATTAAGAACAAAGTAACAGAAATATCTAGGAATACGGCT 426
Qy 361 CTTTGTCTCGAATTAACAAGTTTAGGAGATTCCTTTAGAGGCTATCAACAGTCACTTTGAAGAT 420
Db 427 ATTGCTCTCGAATTAAGAGGCTTAGGAAGAGGCTATAGATCTTTACCAGCAGGCTCTTGAAC 486
Qy 421 TGCTTAGAAAACCGTGTATGATGCAAGAACGAGAAGTGTCTTTTATACCCAAATATATAGCC 480
Db 487 TGCTTAGATAAACCGAATGATGCAAGATCAAGAAGCATTTATCTTGAGCGCTATGTGCT 546
Qy 481 TTAGAACTTTGATTTTCTTAATGCGATGCGCTTTTCGCAATTAGAAAACCAAGAGTTTCCA 540
Db 547 TTAGAACTTTGATTTTCTTAATGCGATGCGCTTTTCGCAATTAGAAAACCAAGAGTTTCCA 606
Qy 541 TTATTAATGATATATGCTCAAGCTGCAAAATTTACACTTATTTATTTAGAGAGTCTCTCT 600
Db 607 TTATTAATGATATATGCTCAAGCTGCAAAATTTACACTTATTTATTTAGAGAGGATGCTCT 666
Qy 601 CTTTGTGTAGTGAATTTTGGGCTTTACATCGCAGAAAATTTCAACGTTTATTTATGAGCGCAA 660
Db 667 CTTTGTGTAGTGAATTTGGGCTTTACATCGCAGAAAATTTCAACGTTTATTTATGAGCGCAA 726
Qy 661 GTGGAACAAACGAGAGATTTATCCGACTATTTGCGTAGAATGCTGATATATACAGTCTTAAT 720
Db 727 ATCAGATATACAGAGGAATATTTCTAACTTGGCTACAAATGGTATATATACAGGGCTAAAT 786
Qy 721 AGCTTCAGAGGGACAAATGCGCAAGTTGGGTGCGTTTATAATCAATTCCTGAGAGATCTA 780
Db 787 AACTTAAGAGGGACAAATGCTGAAGTTGGTGGGTATTAATCAATTCCTGAGAGATCTA 846
Qy 781 ACCTTAGGGGTATTTAGATCTTAGTGGCACTATTTCCCAAGCTATGACACTCGCACTTATCCA 840
Db 847 ACCTTAGGGGTATTTAGATCTTAGTGGCACTATTTCCCAAGCTATGATACTCGCACTTATCCA 906
Qy 841 ATAAATACGAGTCTCAGTTTAAACAGGAAGTTTATACAGACCAATTTGAGCAACAGGG 900
Db 907 ATCAATACGAGTCTCAGTTTAAACAGGAAGTTTATACAGATCCAAATTTGGGAGAACAAAT 966
Qy 901 GTAAAT-ATGCAAGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 954
Db 967 GCACCTCAGGATTTGCAAGTACAAATGGTTTAAATTAATTAATTAATTAATTAATTAAT 1026
Qy 955 ATAGAGACTGCGGTATCCGAAGCCCGCATCTACTTGAATTTCTAGAACAACTTACAAT 1014
Db 1027 ATAGAGCTGCCATTTTCCAGGCTTCCGCTCTACTTGAATTTCTAGAACAACTTACAAT 1086
Qy 1015 TTTAGCACTTCATCAGATGGAGTGTCTATAGGCATATGATTTACTTGGCGGGGACACA 1074
Db 1087 TACAGTGCAATCAAGCCGTTGGAGTAGCACTCAACATATGAATTTATTTGGTGGGACATAGG 1146
Qy 1075 ATTCAATCTCGGCAATAGGAGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1131
Db 1147 CTTAACTTCGCGCCATAGGAGGGACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1206
Qy 1132 ACTTCTAATTAATCCTGTAAAGATTTATCTTCTTCGAGACGTTATTTGGAGCTGAATCA 1191
Db 1207 ACTTCAATTAATCCTGTAAATTTACCTCTCGAGACGTTTATAGAACAGAAATCA 1266
Qy 1192 TATGAGGAGTGTCTTCTATGCGGAAATTTACCTTGAACCTTATTTATGCTGCTCCCTCTGTT 1251
Db 1267 AATGCGAGGACAAATAT-AT-AT-AT-AT-AT-AT-AT-AT-AT-AT-AT-AT-AT-AT-AT 1320
Qy 1252 AGATTTAAATTTTAGGAACCTTCAGAAATATCTTTTGAAGAGGTTACTGCTAACTATAGTCAA 1311
Db 1321 AGATTTAAATTTTAGGAACCTTCAGAAATATTTATGAAGAGGCGCCACTACTTACAGTCAA 1380
Qy 1312 CCTATGAGTCACTCGGCTTCAATTTAAAGATTCAGAAATGAAATTAACCCAGAAACA 1371
Db 1381 CCGTATCAGGAGTTCGGATTTCAATTTATTTGATTCAGAAATGAAATTTACCCAGAAACA 1440
Qy 1372 ACAGAACGACCAATTTATGAATCATATAGTTCATAGGTTATCTCAGATAGGCTCATTTCA 1431
Db 1441 ACAGAACGACCAATTTATGAATCATATAGTTCATAGGTTATCTCATATAGGCTCATTTCA 1500
```

Qy 1432 CAATCTAGGGTGCAATGTAACAGTATATCTTTGGACGACCGTAGTCAGATCGTACAAAT 1491
Db 1501 GGAACACCTTTTGAGAGCACCAGTCTATCTTTGGACGCAATCGTAGTCAGATCGTACGAAT 1560
Qy 1492 ACCATTAGTTCAGATAGCAATAACAACAAATACCAATTTGGTAAATCATTTCAACCTTAATTC 1551
Db 1561 ACGATTGGACCAATAGAAATTACCAAAATACCAATTTGGTAAAGCACTGAATCTTTCAATTC 1620
Qy 1552 GGTAACCTCTGTAGTCAGTGGCCAGGATTTTACAGGAGGGATATAAATCCGAACCTAAACGTT 1611
Db 1621 GGTGTTACTGTGTTGGAGGCGCAGGATTTTACAGGTGGGATATCTTCTGTAGAACAAAT 1680
Qy 1612 AATGGTAGTGTAATAGTATGGGTCCTTAATTTTAAATTAATACATCAATTTACAGCGGTATCGC 1671
Db 1681 ACGGATACATTTGGAGATATACGATTAATTAATTTGCGCATTTATCCCAAAGATATCGC 1740
Qy 1672 GTGAGAGTTCGTTATGCTGCTCTCAACAAATGCTCTGAGGGTAACCTGCGGAGGAGT 1731
Db 1741 GTAAGGATTCGTTATGCTTCTACTACAGATTTTACAATTTTTTACGAGAAATTAATGGAAAC 1800
Qy 1732 ACTACTTTTGATCAAGGATTCCTTAGTACTATGATGCAAAATGAGTCTTTGACATCTCAA 1791
Db 1801 ACTGTTAATTTGTAATTTCTCAAGAACTATGATAGGGGGATTAATTTAGATATAGA 1860
Qy 1792 TCATTTAGATTTGAGAAATTTCTGTAGGTATTTAGTGCATCTGGCAGTCAAACTGCTGGA 1851
Db 1861 AGTTTTAGAACTGCAAGATTTAGTACTCTCTTTTAAATTTTTTAAATGCCAAAGCACATTC 1920
Qy 1852 ATAAGTATAGTAAATATGACGAGTAGACAAACGTTTCACTTTGATAAATTTGAATTCAT 1911
Db 1921 ACATTTGGTCTCGAGATTTTTCAAATCAGGAAGTTTATATAGATAGATCGGAATTTGTT 1980
Qy 1912 CCAATTACTGCAACTTTGGAAGCAGAAATACGATTTAGAAAGGCGCAAGAGGCGTGAAT 1971
Db 1981 CCAGCAGAGTACATTTTGAGGCAGAAATATGATTTAGAAAGACCAAAAGGCGGTGAAT 2040
Qy 1972 GCTCTGTTTACTAATACGAATCCAAAGAGATTTGAAACAGATGTGACAGATTTATCATATT 2031
Db 2041 GCTCTGTTTACTTCTACAAATCCAAAGAGATTTGAAACAGATGTGACAGATTTATCATATT 2100
Qy 2032 GATCAAGTATCCAAATTTAGTGGCGTGTATTCGGATGAATTCGTTAGATGAAAGAGA 2091
Db 2101 GACCAAGTGTCCAAATTTGGTGGCATGTTTATCAGATGAATTTTGTGGATGAGAAGCGA 2160
Qy 2092 GAATTTACTTCGAAAGTGAATATGCGAAACGACTCAGTGAATGAAAGAACTTTACTCCAA 2151
Db 2161 GAATTTTGGAAAGTGAATATGCGAAGGACTCAGTGAATGAAAGAACTTTACTCCAA 2220
Qy 2152 GATCCAACTTCAATCCATCAATTAAGCAACCAAGACTTTCATATCTACTAATGAGCAATCG 2211
Db 2221 GATCCAACTTCAATTCATCAGTGGGCAATTTAAGTTTTCGATCCATCGATGGACAAATCA 2280
Qy 2212 AATTTACATCTATCCATGAACAATCTGAACATGGATGGTGGGAGTGAAGACATTTACA 2271
Db 2281 AACTTCCCTCTAATTAATGAGCTATCTGAACATGGATGGTGGGAGTGGGAATGTTTACC 2340
Qy 2272 ATCCAGGAAGAATGACGTATTTTAAAGAGAAATTTACGTCACTACCGGGGACTTTTAAAT 2331
Db 2341 ATTCAGGAAGGAATGACGTATTTTAAAGAGAAATTTACGTCACTACCGGGTACTTTTAAAT 2400
Qy 2332 GAGTGTATTCGAGCTATTTATATCAAAAAATAGGAGAGTTCGGAATTTAAAGCTTTATAC 2391
Db 2401 GAGTGTATTCGAAATTTATATCAAAAAATAGGAGAGTTCGGAATTTAAAGCTTTATAC 2460
Qy 2392 CGGTACCAATTAAGAGGTATATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGT 2451
Db 2461 CGGTATCAATTAAGAGGTATATTTGAAGATAGTCAAGATCTAGAGATTTTATTTAAATTCGT 2520
Qy 2452 TATTAATCGGAAACATGAAACATTTGGAATGTTCCAGGTACCGAGTCCGATTCGCGCTTCA 2511
Db 2521 TACAAATCGAAGCATGAACATTTGGAATGTTCCAGGTACCGATTCCTTATGGCGCGCTTCA 2580

Qy 2512 GTTGAAGCCCAATCGGAAGGTGGGAGAACCGAATCGATGCGCACCAATTTTGAATGG 2571
Db 2581 GTTGAAGCCCAATCGGAAGGTGGGAGAACCGAATCGATGCGCACCAATTTTGAATGG 2640
Qy 2572 AATCCTGATCTAGATTTGTTCTCGAGAGTGGAGAAAAATGTGCGCATCTATCCCATCAT 2631
Db 2641 AATCCTGATCTAGATTTGTTCTCGAGAGTGGAGAAAAATGTGCGCATCTATCCCATCAT 2700
Qy 2632 TTCTCTTTGGATTTGATTTGGATGCACAGACTTTCATCAGAAATCTAGCGCTGTGGGTG 2691
Db 2701 TTCACTTTGGATTTGATTTGGATGCACAGACTTTCATCAGAACTTAGCGCTGTGGGTG 2760
Qy 2692 GTATTCAAGATTTAAGACGAGGAGGTCATGCAAGACTAGGGAATCTGGAAATTTTATGAA 2751
Db 2761 GTATTCAAGATTTAAGACGAGGAGGTCATGCAAGACTTAGGAAATCTGGAAATTTTATGAA 2820
Qy 2752 GAGAAACATTTATTAGGAGAGCACTGTCTCGTGTGAAGAGAGCAGAGAAAAAATGAGA 2811
Db 2821 GAGAAACATTTATTGGAGAGCACTGTCTCGTGTGAAGAGAGCGGAAAAAATGAGA 2880
Qy 2812 GACAAACGCTGAAAACTTACAAATTTGAAAAACAAACGAGTATATACAGAGGCAAAAGAACT 2871
Db 2881 GACAAACGCGAAAACTTACAAATTTGAAAAACAAACGAGTATATACAGAGGCAAAAGAACT 2940
Qy 2872 GTGGATGCTTTATTTGTAGATTTCTCAATATAATAGATTAACAAGCGGATACAACATTTGGC 2931
Db 2941 GTGGATGCTTTATTTGTAGATTTCTCAATATATGATCAATTTACAAGCGGATACAACATTTGGC 3000
Qy 2932 ATGATTCATGCGCGCAGATAAATCTGTTCAATCGAAATTCGAGAGGCTTATCTGTCAGAAATTA 2991
Db 3001 ATGATTCATGCGCGCAGATAAATCTGTTCAATCGAAATTCGAGAGGCTTATCTTTCAGAAATTA 3060
Qy 2992 TCTGTTATCCCGGTGTAAATGCGGAAATTTTTCAGAAATTTAGAAAGTGCATTTATCACT 3051
Db 3061 CCGTGTATCCAGGTGTAAATGCGGAAATTTTTCAGAAATTTAGAAAGTGCATTTATCACT 3120
Qy 3052 GCAATCTCCCTATACATGCGAGAAATGTCTGTTTAAAAATGGTGATTTTAAATTAATGAAATTA 3111
Db 3121 GCAATGCTTTATACATGCGAGAAATGTCTGTTTAAAAATGGTGATTTTAAATTAATGAAATTA 3180
Qy 3112 GCATGCTGGAATGTAAAGGGCATGTAGATGTACAACAGAGCCATCAACGTTCTGTCCTT 3171
Db 3181 ACATGTTGGAATGTAAAGGGCATGTAGATGTACAACAGAGCCATCACTGTTCTGACCTT 3240
Qy 3172 GTTATCCGAAATGGGAAGCAGAAAGTGTCAACAGCAGTTCGCGTCTCTCGGGGCGTGGC 3231
Db 3241 GTTATCCGAAATGGGAAGCAGAAAGTGTCAACAGCAGTTCGCGTCTCTCGGGGCGTGGC 3300
Qy 3232 TATATCTCCGTGTACAGGTTACAAAGAGGATATGGAAGGGTGTGTGTAAACGATCCAT 3291
Db 3301 TATATCTCCGTGTACAGGTTACAAAGAGGATATGGAAGGGTGTGTGTAAACGATCCAT 3360
Qy 3292 GAAATCGGAACAAATACAGAGCACTAAAATTTTAAAACTGTGAAGAGAGGAGTGTAT 3351
Db 3361 GAAATCGGAACAAATACAGAGCACTAAAATTTTAAAACTGTGAAGAGAGGAGTGTAT 3420
Qy 3352 CCAACGGATACAGAACTGTGAATTAATTAATCTGCACACCAAGGTAACAGAGTATGTAAT 3411
Db 3421 CCAACGGATACAGAACTGTGAATTAATTAATCTGCACACCAAGGTAACAGAGTATGTAAT 3480
Qy 3412 TCCCGTAAATGCTGGATATGAGATATGATATGAAAGTTGATATCTACAGCATCTGTTAAATAC 3471
Db 3481 TCCCGTAAATGCTGGATATGAGATATGATATGAAAGTTGATATCTACAGCATCTGTTAAATAC 3540
Qy 3472 AAACCGACTTTTCAAGAGAAACGTATACAGATGTACGAGAGATTAATCTGTCGATAT 3531
Db 3541 AAACCGACTTTTCAAGAGAAACGTATACAGATGTACGAGAGATTAATCTGTCGATAT 3600
Qy 3532 GACAGAGGGTATGTGAATTAATTCACCACTACAGCTGGTGTATATGACAAAGAAATTAGAA 3591
Db 3601 GACAGAGGGTATGTGAATTAATTCACCACTACAGCTGGTGTATATGACAAAGAAATTAGAA 3660
Qy 3592 TACTTCCAGAAACCGATAGGATGGAATGGAATTTGGAATTTGGAAGAAACGGAAGGAAATTTATT 3651

Db	3661	TACTTCCAGAAACAGATACAGTATGATTCGAGAAACGGAAGGAAAGTTTATT	3720
Qy	3652	GTACACAGCGTGAATTAATCTCTTATGAGGAATAG	3687
Db	3721	GTAGTAGCGTGAAGTACTCTCTCATGGAAGATAG	3756
RESULT 15			
LOCUS	I38760	3934 bp	DNA
DEFINITION	Sequence 3 from patent US 5616319.		
ACCESSION	I38760		
VERSION	I38760.1	GI:2083238	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 3934)		
AUTHORS	Donovan, W.P., Tan, Y., Jany, C.S. and Gonz alez, J.E.M. Jr.		
TITLE	Bacillus thuringiensis cryET5 gene and related plasmids, bacteria and insecticides		
JOURNAL	Patent: US 5616319-A 3 01-APR-1997;		
FEATURES	Location/Qualifiers		
source	1..3934		
ORIGIN	/organism="unknown"		
	/mol_type="unassigned DNA"		
Query Match	72.9%;	Score 2687.2;	DB 6; Length 3934;
Best Local Similarity	83.5%;	Pred. No. 0;	
Matches 3088;	Conservative	0; Mismatches 593;	Indels 15; Gaps 3;
Qy	1	TTGACTTCAATAGAGAAATGAGAAATGAAATTAATAATGCTTTATCGATTCGAGCTGTA	60
Db	67	TTGACTTCAATAGAGAAATGAGAAATGAAATTAATAATGCTTTATCGATTCGAGCTGTA	126
Qy	61	TCGAATCAATCCACACAAATGATCTATCACCAGATGCTCGTATGAGGATTCCTTCTGT	120
Db	127	TCGAATCCTTCCAGCAATGAATCTATCACCAGATGCTCGTATGAGGATTCCTTCTGT	186
Qy	121	ATAGCCGAGGGGAATATATCAATCCACTTGTTTAGCGCATCAACAGTCCAAAACGGGTATT	180
Db	187	GTAGCCGAGGTGAACAATATGATCCATTTGTTAGCGCATCAACAGTCCAAAACGGGTATA	246
Qy	181	AACATAGCTGTAGAATACTAGGTGATATAGGCGTACCGTTTGTCTGACAAATAGTAGT	240
Db	247	AACATAGCTGTAGAATACTAGGCGTATTTGGGCGTATTTAGTGTGCGTTTGTCTGACAACTAGTAGT	306
Qy	241	TTTTATAGTTTTCTTTGTTGGAATTAATGCGCCGCGAGAGATCAGTGGGAAATTTTC	300
Db	307	TTTTATAGTTTTCTTTGTTGGGAATTTATGGCTAGTGGCAGATCCATGGGAAATTTTC	366
Qy	301	CTAGAACATGTGGAACAACTTATAAATCAACAAATACAGAAATGCTAGGAATACGGCA	360
Db	367	CTGGAACATGTAGAACAACTTATAGAACAAACAGTAAACAGAAATACTAGGAATACGGCT	426
Qy	361	CTTCTCGATTTACAAGGTTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT	420
Db	427	ATTCTCGATTTAGAGGTTTAGGAGGCTATAGATCTTACGAGGCTCTTGAAGT	486
Qy	421	TGGCTAGAAAACCGTGATGATGCAAGAACGAGAGTGTCTTTTATACCCAAATATATAGCC	480
Db	487	TGGTTAGATTAACCGAAATGATGCAAGATCAAGAGAGCATTAATCTTGAGCGCTATGTTGCT	546
Qy	481	TTGAACCTTGATTTCTTAATGCGATGCGCTTTTCGCAATTAGAACACCAAGAGATTCCA	540
Db	547	TTGAACCTTGACATTACTACTGCTATACCGCTTTTCAGATACGAATGAGAGATTCCA	606
Qy	541	TTATTAATGATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGCTCT	600
Db	607	TTATTAATGATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGCTCT	666

Db	1741	GTAAGGATTCGGTATATGCTTCTACTACAGATTTTACAATTTTTTTCACGAGAAATTAATGGAACC	1800
Qy	1732	ACTACTTTTGATCAAGGATTCCTTAGTACTATAGTGCAAATGAGTCTTTTGACATCTCAA	1791
Db	1801	ACTGTTAATATGGTAAATTTCTCAAGAACTATGAAATAGGGGGGTAATATTTAGAAATATAGA	1860
Qy	1792	TCATTTAGATTTGCAGAAATTTCTCTGTAGTATTTAGTGCATCTGGCAGTCAAACTGCTGGGA	1851
Db	1861	AGTTTTAGAACTGCAGGATTTAGTACTCTTTTAAATTTTTTAAATGCCAAAGCACATTC	1920
Qy	1852	ATAGTATATAGTAATATATGCAGGTAGACAACGGTTTCACTTTGATATAAAATTTGAATTCATT	1911
Db	1921	ACATTTGGGTGCTCAGATGTTTTTCCAAATCAGGAAGTTTATATAGATAGAGTCGAATTTGTT	1980
Qy	1912	CCAATTACTCAACCTTCGAAGCAGAAATACGATTTAGAAAGGCGCAAGAGGCGGTGAAT	1971
Db	1981	CCAGCAGGTATACATTTGAGCGAATATGATTTAGAAAGACACAAAGGCGGTGAAT	2040
Qy	1972	GCTCTGTTTACTAATACGAATCCAAGAAGATTTGAAAAAGATGTGACAGATTTATCATATT	2031
Db	2041	GCTCTGTTTACTCTACAAATCCAAGAAGATTGAAAACAGATGTGACAGATTTATCATATT	2100
Qy	2032	GATCAAGTATCCAAATTTAGTGGCGGTGTTATTCGGATGAATTTCTGTTTATGATGAAAAGAGA	2091
Db	2101	GACCAAGTCCAAATATGTTGGTGGCATGTTTATCAGATGAAATTTTCTTGGATGAGAAGCGA	2160
Qy	2092	GAATTACTTCGAGAAAGTGAATATGCCAAAGCATCTAGTCATGATGAAGAAACTTTACTCCAA	2151
Db	2161	GAATTTATTCGAGAAAGTGAATATGCCAAGCGACTCAGTGATGAAGAAACTTTACTCCAA	2220
Qy	2152	GATCCAACTTCACATCCATCAATTAAGCAACACAGACTTTTCATATCTATAATAGCAATCG	2211
Db	2221	GAICCAAACTTCACATTCACAGTGGGCAATTAAGTTTTCGCATCCATCGATGGACAATCA	2280
Qy	2212	AATTTACATCTATCCATGAACAACTCTGAACTAGATGGTGGGGAAGTGAAGACAATTACA	2271
Db	2281	AACTTCCCTCTATTAAATGAGCTATCTTGAACTAGTATGTTGGGGAAGTGGGAATTTTACC	2340
Qy	2272	ATCCAGGAAGGAATGACGTATTTTAAAGAAATTTACGTTCACACTACCGGGGACTTTTAAAT	2331
Db	2341	ATTGAGGAAGGAATGACGTATTTTAAAGAAATTTACGTTCACACTACCGGGTACTTTTAAAT	2400
Qy	2332	GAGTGTATTCCGACGTATTTATATCAAAAAATAGCAGAGTCGGAATTTAAAGCTTTATACT	2391
Db	2401	GAGTGTATTCCAAATTTATTTATCAAAAAATAGGAGAGTCAGAA'TTAAAGCTTTATACG	2460
Qy	2392	CGGTACCAATTAAGAGGGTATATTGAAGATAGTCAAGATTTAGAGATPATATTTGATTCGT	2451
Db	2461	CGGTATCAATTAAGAGGGTATATTGAAGATAGTCAAGATCTAGAGATTTTATTTAAATTCGT	2520
Qy	2452	TATTAATCGGAAACATGAAACATTCGATGTTCCAGGTACCGAGTCCGTATCGCGCTTTCA	2511
Db	2521	TACAATGCAAGCATGAAACATTCGATGTTCTCAGGTACCGAATTTCCCTATTTGGCGCTTTCA	2580
Qy	2512	GTTGAAAGCCAAATCGGAAGGTGCGGAGAACCCGAATCGATCGGCACACACATTTTGAATGG	2571
Db	2581	GTTGAAAGCCAAATCGGAAGGTGCGGAGAACCAATCGATCGGCACACACATTTTGAATGG	2640
Qy	2572	AATCCTGATCTAGATGTTTCTGAGAGATGGAGAAAAATTTGTGCGCATCATTTCCCATCAT	2631
Db	2641	AATCCTGATCTAGATGTTTCTGAGAGATGGAGAAAAATTTGTGCGCATCATTTCCCATCAT	2700
Qy	2632	TTCTCTTTGGATATTGATATTTGGATGCACAGACTTGGCATGAGATCTAGGCGTGTGGGTG	2691
Db	2701	TTCACTTTGGATATTGATTTGGGTGCACAGACTTGCATGAGAACTTAGCGCTGTGGGTG	2760
Qy	2692	GTATTTCAAGATTTAAGACCGCAGGAAGTTCATGCAAGACTTAGGGAATCTGGAATTTATTGAA	2751
Db	2761	GTATTTCAAGATTTAAGACCGCAGGAAGTATTGCAAGATTTAGGAATCTGGAATTTATCGAA	2820
Qy	2752	GAGAAACCATTTATTAGGAGAAGCACTGCTCTGTTGTAAGAGACAGAGAAAAATTTGGAGA	2811

Db	2821	GAGAAACCA	TTAA	TTGGAGAG	CAC	TGTCT	GTGTGAAGAG	CGG	AAAA	AAAA	TTGGAGA	2886
Qy	2812	GACAAACGT	GAAAACT	CAAA	TTGGAA	CAAA	ACG	AGTATATACAGAGG	CAAA	AGAA	AGAGCT	2871
Db	2881	GACAAACGG	AAAACT	CAAA	TTGGAA	CAAA	ACG	AGTATATACAGAGG	CAAA	AGAA	AGAGCT	2940
Qy	2872	GTGGATCG	TTTTAT	TTGTAG	ATCT	CAAT	ATAATAG	ATTACAAGCGG	ATACAACAT	TTGGC	2931	
Db	2941	GTGGATCG	TTTTAT	TTGTAG	ATCT	CAAT	ATAATAG	ATTACAACG	ATACAACAT	TTGGC	3000	
Qy	2932	ATGATTCAT	CGCGCAG	ATAAC	TTGTT	TCAT	CGAATTCGAGAGG	CTTATCTGT	CAGAA	TTA	2991	
Db	3001	ATGATTCAT	CGCGCAG	ATAAC	TTGTT	TCAT	CGAATTCGAGAGG	CTTATCTGT	CAGAA	TTA	3060	
Qy	2992	TCTGTAT	CCCGGTG	TAAAT	TGCGAA	ATTTTT	TGAAGAA	TTAGAA	GGTCGCAT	TATCACT	3051	
Db	3061	CCTGTAT	CCCGGTG	TAAAT	TGCGAA	ATTTTT	TGAAGAA	TTAGAA	GGTCGCAT	TATCACT	3120	
Qy	3052	GCAATCT	CCCTATA	CGAT	TGCGAGAA	ATGTCGT	TAAAAAT	GGTGA	TTTTAA	TAAATGGA	3111	
Db	3121	GCAATGCT	TTATAC	GATGCGAGAA	ATGTCGT	TAAAAAT	GGTGA	TTTTAA	TAAATGGA	3180		
Qy	3112	GCATCTG	GGAATG	TAAAGG	CGATG	TACACAGAG	CCCAT	CCCGTT	CTGTCT	3171		
Db	3181	ACATGTT	GGAATG	TAAAGG	CGATG	TACACAGAG	CCCAT	CCCGTT	CTGTCT	3240		
Qy	3172	GTATCC	CAAGATGG	GAAGCAG	AGTGT	CAACAG	CGTCCG	CTCTCGGG	CGTGGC	3231		
Db	3241	GTATCC	CAAGATGG	GAAGCAG	AGTGT	CAACAG	CGTCCG	CTCTCGGG	CGTGGC	3300		
Qy	3232	TATATCCT	CCGTGT	CACAGCGT	CAAA	AGGGATATGGAGAG	GGTGTGT	TAAACG	ATCCAT	3291		
Db	3301	TATATCCT	CCGTGT	CACAGCGT	CAAA	AGGGATATGGAGAG	GGTGTGT	TAAACG	ATCCAT	3360		
Qy	3292	GAAATCG	AGAACAT	ATACAGAG	CAAT	TAATTTAA	AACTGTG	GAAGAGG	AGAGTGT	3351		
Db	3361	GAAATCG	AGAACAT	ATACAGAG	CAAT	TAATTTAA	AACTGTG	GAAGAGG	AGAGTGT	3420		
Qy	3352	CCAACG	GATACAG	ACGTGT	ATATGATTTAT	CTGCAC	CAACAGGT	CACAGAG	TATGT	3411		
Db	3421	CCAACG	GATACAG	ACGTGT	ATATGATTTAT	CTGCAC	CAACAGGT	CACAGAG	TATGT	3480		
Qy	3412	TCCCGT	TAATGCTG	GATATGAG	GTGCA	TATGAAG	TTGATCTACAG	CATCTG	TTAAT	3471		
Db	3481	TCCCGT	TAATGCTG	GATATGAG	GTGCA	TATGAAG	TTGATCTACAG	CATCTG	TTAAT	3540		
Qy	3472	AAACCG	ACTTATG	AGAAAGAA	ACGTATACAG	ATGTTACG	AGAGATAT	CAATGT	GTGAAT	3531		
Db	3541	AAACCG	ACTTATG	AGAAAGAA	ACGTATACAG	ATGTTACG	AGAGATAT	CAATGT	GTGAAT	3600		
Qy	3532	GACAG	GGGTATGT	GAAATAT	CCACCACT	ACCAGCTG	TTATATG	ACAAAG	AGAT	3591		
Db	3601	GACAG	GGGTATGT	GAAATAT	CCACCACT	ACCAGCTG	TTATATG	ACAAAG	AGAT	3660		
Qy	3592	TACTT	CCAGAA	CCGAT	TAAGTAT	TGAGAT	TTGAGAT	TTGAGAA	ACGAA	3651		
Db	3661	TACTT	CCAGAA	CCGAT	TAAGTAT	TGAGAT	TTGAGAT	TTGAGAA	ACGAA	3720		
Qy	3652	GTAG	ACGCTG	GGAAT	TACTCT	TTATGG	AGGAAT	ATAG	3687			
Db	3721	GTAG	ATAGCT	GGAAT	TACTCT	TTATGG	AGGAAT	ATAG	3756			

Search completed: April 21, 2005, 00:32:35
Job time : 15279.secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2005, 16:05:50 ; Search time 1752 Seconds
(without alignments)
12457.812 Million cell updates/sec

Title: US-10-614-524-1
Perfect score: 3687
Sequence: 1 ttgactcaaatggaagaaa.....tactcttatggagggaatag 3687

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3687	100.0	3687	AAH28240	AAH28240 Nucleotid
2	3666.2	99.4	4173	AAH28240	Aas02482 B. thurin
3	3266.6	88.6	3684	AAV15222	AAV15222 Bacillus
4	3266.6	88.6	3684	AAH92221	AAH92221 Bacillus
5	3260.2	88.4	3684	AAH92221	Aas02489 B. thurin
6	3132	84.9	4074	AAQ56804	AAQ56804 Bacillus
7	2687.2	72.9	3932	AAQ64112	AAQ64112 cryET5 ge
8	2687.2	72.9	3932	AAQ09160	Aas09160 B. thurin
9	2687.2	72.9	3934	AAT95051	Aat95051 DNA encod
10	2687.2	72.9	3934	AAT68434	Aat68434 CryET5 ge
11	2687.2	72.9	3934	AAV83927	AAV83927 DNA encod
12	2684.2	72.8	3687	ADK98478	Adk98478 B thuring
13	2630.2	71.3	3627	AAQ04854	Aas04854 Bacillus
14	2541.4	69.0	3663	AAAS00422	Aas00422 B. thurin
15	2541.2	68.9	3663	AAAS00421	Aas00421 B. thurin
16	2394.2	64.7	3624	AAAS04853	Aas04853 Bacillus
17	2004.8	54.4	3651	AAAS04852	Aas04852 Bacillus
18	1793.4	48.6	3522	AAT16558	Aat16558 B.t. toxi
19	1791.8	48.6	3522	AAQ06830	AAQ06830 Sequence
20	1791.8	48.6	3522	AAT49592	Aat49592 Bacillus

21	1790.2	48.6	3522	2	AAQ26928	AAQ26928 DNA encod
22	1662	45.1	4106	2	AAV16214	AAV16214 DNA encod
23	1662	45.1	4106	2	AAV09318	AAV09318 Bacillus
24	1581.6	42.9	3558	2	AAQ60619	AAQ60619 Polynucle
25	1501	40.7	3766	2	AAQ77858	AAQ77858 Crystal p
26	1438.6	39.0	8854	2	AAZ20086	AAZ20086 DNA encod
27	1432.4	38.9	3713	2	AAQ64111	AAQ64111 cryET4 ge
28	1432.4	38.9	3713	2	AAT95050	Aat95050 DNA encod
29	1432.4	38.9	3713	2	AAT68435	Aat68435 CryET4 ge
30	1432.4	38.9	3713	2	AAV83926	AAV83926 DNA encod
31	1432.4	38.9	3713	2	AAZ09159	Aaz09159 B. thurin
32	1411.4	38.3	4020	2	AAQ14648	AAQ14648 cryLF gen
33	1408.4	38.2	3522	2	AAQ10182	AAQ10182 Lepidopte
34	1408.4	38.2	3522	2	AAQ47291	AAQ47291 Delta end
35	1408.4	38.2	3522	2	AAV62081	AAV62081 Plasmid p
36	1408.4	38.2	3522	2	AAQ60606	AAQ60606 wild-type
37	1408.4	38.2	3522	4	AAQ00565	Aas00565 B. thurin
38	1406.8	38.2	3522	2	AAT18722	Aat18722 Codon-rew
39	1405.2	38.1	3522	2	AAT05270	Aat05270 CryIF tox
40	1402.2	38.0	5170	12	ADK98488	Adk98488 B thuring
41	1402.2	38.0	5600	12	ADK98490	Adk98490 B thuring
42	1402.2	38.0	6600	12	ADK98482	Adk98482 B thuring
43	1402.2	38.0	7000	12	ADK98485	Adk98485 B thuring
44	1402	38.0	3534	2	AAV31162	AAV31162 Bacillus
45	1402	38.0	3534	6	ABK11833	ABK11833 DNA encod

ALIGNMENTS

RESULT 1
AAH28240
ID AAH28240 standard; DNA; 3687 BP.
XX
AC AAH28240;
XX
DT 05-SEP-2001 (first entry)
XX
DE Nucleotide sequence of a CryIbF insecticidal protein.
XX
KW CryIbF; insecticidal protein; CryIb; Cry9Fa; lepidoptera; coleoptera;
KW insect pest; transgenic plant; ss.
XX
OS Bacillus thuringiensis.
XX
FH Key Location/Qualifiers
FT CDS 1..3687
FT /tag= a
FT /product= "CryIbF insecticidal protein"
XX
XX WO200147952-A2.
XX
PD 05-JUL-2001.
XX
PF 19-DEC-2000; 2000WO-EP013184.
XX
PR 28-DEC-1999; 99US-0173387P.
XX
PA (AVET) AVENTIS CROPS SCIENCE NV.
XX
XX Arnaut G, Boets A, Damme N, Mathieu E, Vanneste S, Van Rie J;
XX WPI: 2001-425619/45.
XX P-PSDB; AAB84628.
XX
XX Novel insecticidal proteins CryIb, Cry9Fa and CryIbF, derived from
XX Bacillus thuringiensis, useful for controlling insects in plants.
XX Claim 5; Page 30-36; 65pp; English.
XX
XX The present sequence encodes CryIbF, an insecticidal protein derived from
XX Bacillus thuringiensis. The specification also describes CryIb and
XX Cry9Fa. The Cry proteins have activity against lepidopteran and

CC coleopteran insect pests. Cry1Bf, CryIId and Cry9Fa polynucleotides and
CC polypeptides are useful for obtaining a plant with resistance to insects.
CC Cry polynucleotides are useful for producing transgenic plants which are
CC resistant to insects

XX
SQ Sequence 3687 BP; 1219 A; 640 C; 805 G; 1023 T; 0 U; 0 Other;
Query Match 100.0%; Score 3687; DB 4; Length 3687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	TTGACCTCAAAATAGGAAATAGAAATGAAATATATAATGCTTTATCGATTCAGCTGTA	60
Db	1	TTGACCTCAAAATAGGAAATAGAAATGAAATATATAATGCTTTATCGATTCAGCTGTA	60
Qy	61	TCCGAATCATTTCCACACAAATGGATCTATCACAGATGCTCGTATGAGGATTCCTTTGTG	120
Db	61	TCCGAATCATTTCCACACAAATGGATCTATCACAGATGCTCGTATGAGGATTCCTTTGTG	120
Qy	121	ATAGCCGAGGGGAATATATCAATCCACTTGTAGCGCATCAACAGTCCAAACGGGTATT	180
Db	121	ATAGCCGAGGGGAATATATCAATCCACTTGTAGCGCATCAACAGTCCAAACGGGTATT	180
Qy	181	AACATAGCTGGTAGAATATAGTGTATATAGGGGTACCGTTTCTGGACAAATAGCTAGT	240
Db	181	AACATAGCTGGTAGAATATAGTGTATATAGGGGTACCGTTTCTGGACAAATAGCTAGT	240
Qy	241	TTTTATAGTTTTCTTTGTGTGTAATATATGCGCGGGGAGAGATCAGTGGGAAATTTTC	300
Db	241	TTTTATAGTTTTCTTTGTGTGTAATATATGCGCGGGGAGAGATCAGTGGGAAATTTTC	300
Qy	301	CTAGAACATGTCGAAACAACTTATAAATCAACAAATAACAGAAATGCTAGGAATACGGCA	360
Db	301	CTAGAACATGTCGAAACAACTTATAAATCAACAAATAACAGAAATGCTAGGAATACGGCA	360
Qy	361	CTTGCTCGATTACAGTTTATAGGATTCCTTTAGAGCCCTATCAACAGTCACTTGAAGAT	420
Db	361	CTTGCTCGATTACAGTTTATAGGATTCCTTTAGAGCCCTATCAACAGTCACTTGAAGAT	420
Qy	421	TGGCTAGAAACCGTGATGATGCAAGAACGAGAGTGTCTTTATACCCCAATATATAGCC	480
Db	421	TGGCTAGAAACCGTGATGATGCAAGAACGAGAGTGTCTTTATACCCCAATATATAGCC	480
Qy	481	TTAGAACTTGATTTTCTTAATGCGATGCCCTTTTCGCAATTAGAACCAAGAGTTCCA	540
Db	481	TTAGAACTTGATTTTCTTAATGCGATGCCCTTTTCGCAATTAGAACCAAGAGTTCCA	540
Qy	541	TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTTATTTGAGAGATGCCCT	600
Db	541	TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTTATTTGAGAGATGCCCT	600
Qy	601	CTTTTGGTATGTAATTTGGGCTTACATCCAGGAAATTCAGTTTATTTATGAGCGCAA	660
Db	601	CTTTTGGTATGTAATTTGGGCTTACATCCAGGAAATTCAGTTTATTTATGAGCGCAA	660
Qy	661	GTGGAAACAAACGAGAGATTTCCGACTATTGGTGTAGAAATGATATACAGCTCTAAAT	720
Db	661	GTGGAAACAAACGAGAGATTTCCGACTATTGGTGTAGAAATGATATACAGCTCTAAAT	720
Qy	721	AGCTTGAGAGGGAACAAATGCGCAAGTTGGTGGCTTATTAATCAATTCCTGAGATCTA	780
Db	721	AGCTTGAGAGGGAACAAATGCGCAAGTTGGTGGCTTATTAATCAATTCCTGAGATCTA	780
Qy	781	ACGTTAGGGGTATAGATCTAGTGGCACTATTCGAGCTATGACACTGCCACTTATCCA	840
Db	781	ACGTTAGGGGTATAGATCTAGTGGCACTATTCGAGCTATGACACTGCCACTTATCCA	840
Qy	841	ATAAATACGAGTGTCTAGTTTAAACAGGGAAGTTTATACAGACGCAATTTGAGCAACAGGG	900
Db	841	ATAAATACGAGTGTCTAGTTTAAACAGGGAAGTTTATACAGACGCAATTTGAGCAACAGGG	900
Qy	901	GTAATAATGCCAAGTATGAAATGGTATATAATAATGCACTTCCTGTTTCCGCTATAGAG	960
Db	901	GTAATAATGCCAAGTATGAAATGGTATATAATAATGCACTTCCTGTTTCCGCTATAGAG	960

Db	901	GTAATAATGCCAAGTATGAAATGGTATATAATAATGCACTTCCTGTTTCCGCTATAGAG	960
Qy	961	ACTGCGGTTATCCGAAGCCGCAATCTACTTGATTTCTAGAACAACTTACAAATTTTATAGC	1020
Db	961	ACTGCGGTTATCCGAAGCCGCAATCTACTTGATTTCTAGAACAACTTACAAATTTTATAGC	1020
Qy	1021	ACTTCAATCAAGTGGAGTGTCTACTAGGCATATAGCTTACTTGGCGGGGACACAAATTCAA	1080
Db	1021	ACTTCAATCAAGTGGAGTGTCTACTAGGCATATAGCTTACTTGGCGGGGACACAAATTCAA	1080
Qy	1081	TCTCGGCCAATAGGAGCGGATTAATATACCTCAACCATGGGTCTACCAATCTCTATT	1140
Db	1081	TCTCGGCCAATAGGAGCGGATTAATATACCTCAACCATGGGTCTACCAATCTCTATT	1140
Qy	1141	AATCCTGTAGATTTATCATTTCTCTCGAGAGCTATATTTGAGCTCAATCATATGACGGA	1200
Db	1141	AATCCTGTAGATTTATCATTTCTCTCGAGAGCTATATTTGAGCTCAATCATATGACGGA	1200
Qy	1201	GTGCTTCTATGGGAAATTTACCTTGAACTTATTCATGGTGTCCCTACTGTTAGATTTAAT	1260
Db	1201	GTGCTTCTATGGGAAATTTACCTTGAACTTATTCATGGTGTCCCTACTGTTAGATTTAAT	1260
Qy	1261	TTTAGGAACTCTCAGAAATCTTTTGAAGAGTCTGCTAACTATATAGTCAACCTATGAG	1320
Db	1261	TTTAGGAACTCTCAGAAATCTTTTGAAGAGTCTGCTAACTATATAGTCAACCTATGAG	1320
Qy	1321	TCACCTGGGCTTCAATTTAAAGAGTTCAGAACTGAAATTTACCAACAGAAACACAGAACGA	1380
Db	1321	TCACCTGGGCTTCAATTTAAAGAGTTCAGAACTGAAATTTACCAACAGAAACACAGAACGA	1380
Qy	1381	CCAAATTTATCAATCATATAGTTCATAGTTTATCTCACATAGGGCTCATTTTCAATCTAGG	1440
Db	1381	CCAAATTTATCAATCATATAGTTCATAGTTTATCTCACATAGGGCTCATTTTCAATCTAGG	1440
Qy	1441	GTGCTATGACAGTATTTCTTGGACGCACTGAGTGCAGATCGTCAAAATACCATTAGT	1500
Db	1441	GTGCTATGACAGTATTTCTTGGACGCACTGAGTGCAGATCGTCAAAATACCATTAGT	1500
Qy	1501	TCAGATAGCATTAACAAATACCAATTTGTAATAATCAATCAACCTTAATTCAGGTACTCT	1560
Db	1501	TCAGATAGCATTAACAAATACCAATTTGTAATAATCAATCAACCTTAATTCAGGTACTCT	1560
Qy	1561	GTAGTCAAGTGGCCAGGATTTTACAGAGGGATATTAATCCGAACTAACTGTTAATGGTAGT	1620
Db	1561	GTAGTCAAGTGGCCAGGATTTTACAGAGGGATATTAATCCGAACTAACTGTTAATGGTAGT	1620
Qy	1621	GTACTAAGTATGGTCTTAAATTTTAAATATACATCAATTAACAGCGGTATCGGTGAGATT	1680
Db	1621	GTACTAAGTATGGTCTTAAATTTTAAATATACATCAATTAACAGCGGTATCGGTGAGATT	1680
Qy	1681	CGTTATGCTGCTTCTCAAAACAAATGGTCTGAGGGTAACTGTCCGAGGGAGTACTACTTTT	1740
Db	1681	CGTTATGCTGCTTCTCAAAACAAATGGTCTGAGGGTAACTGTCCGAGGGAGTACTACTTTT	1740
Qy	1741	GATCAAGGATTCCTTAGTACTATGAGTCAAAATGAGTCTTTTGACATCTCAATCATTTAGA	1800
Db	1741	GATCAAGGATTCCTTAGTACTATGAGTCAAAATGAGTCTTTTGACATCTCAATCATTTAGA	1800
Qy	1801	TTTGAGAAATTTCTCTAGTATAGTGCATCTGGCAGTCAAACTGCTGGGAATAGTATA	1860
Db	1801	TTTGAGAAATTTCTCTAGTATAGTGCATCTGGCAGTCAAACTGCTGGGAATAGTATA	1860
Qy	1861	AGTAATAATGAGGTAGACAAACGTTTTCATTTTGAATAAATTCATTTCCAAATTACT	1920
Db	1861	AGTAATAATGAGGTAGACAAACGTTTTCATTTTGAATAAATTCATTTCCAAATTACT	1920
Qy	1921	GCAACCTCTCGAAGCAGAAATACGATTTTGAAGGGCCAAAGAGCGGTGAATGCTCTGTTT	1980
Db	1921	GCAACCTCTCGAAGCAGAAATACGATTTTGAAGGGCCAAAGAGCGGTGAATGCTCTGTTT	1980
Qy	1981	ACTAATAACGAAATCCAGAGATTTGAAACAGATGTCAGAGATTTATCATTTGATCAAGTA	2040
Db	1981	ACTAATAACGAAATCCAGAGATTTGAAACAGATGTCAGAGATTTATCATTTGATCAAGTA	2040

```
Qy 2041 TCCAATTAGTGGCGTGTATCGATCAATTTCTGCTTAGATGAAGAGAGAAATTAATT 2100
Db 2041 TCCAATTAGTGGCGTGTATCGATCAATTTCTGCTTAGATGAAGAGAGAAATTAATT 2100
Qy 2101 GAGAAAGTGAATATGCGAAACGACTCAGTGAATGAAGAAATTAATTCTTCCAAAGATCCAAAC 2160
Db 2101 GAGAAAGTGAATATGCGAAACGACTCAGTGAATGAAGAAATTAATTCTTCCAAAGATCCAAAC 2160
Qy 2161 TTCAATCCATCAATAAGCAACAGACTTCAATCTACTAATGAGCAATCGAATTTTCA 2220
Db 2161 TTCAATCCATCAATAAGCAACAGACTTCAATCTACTAATGAGCAATCGAATTTTCA 2220
Qy 2221 TCTATCCATGACAACTCTGACATGGATGGTGGGGAAGTGNACATTAACATCCAGAA 2280
Db 2221 TCTATCCATGACAACTCTGACATGGATGGTGGGGAAGTGNACATTAACATCCAGAA 2280
Qy 2281 GGAATGACAGTATTTAAAGAGAAATTCGTCACTACCGGGGACTTTTAATGAGTGTAT 2340
Db 2281 GGAATGACAGTATTTAAAGAGAAATTCGTCACTACCGGGGACTTTTAATGAGTGTAT 2340
Qy 2341 CCGACGATTTATCAAAAATAGGAGAGTGGAAATTAAGGCTTATCTCGTACCAA 2400
Db 2341 CCGACGATTTATCAAAAATAGGAGAGTGGAAATTAAGGCTTATCTCGTACCAA 2400
Qy 2401 TTAAGAGGTATATTAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAATGCG 2460
Db 2401 TTAAGAGGTATATTAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAATGCG 2460
Qy 2461 AATCATGAACATTTGGATGTTCCAGGTACCGAGTCCGTATGGCCGCTTTCAGTTGAAAGC 2520
Db 2461 AATCATGAACATTTGGATGTTCCAGGTACCGAGTCCGTATGGCCGCTTTCAGTTGAAAGC 2520
Qy 2521 CCAATCGGAAGTCCGGAGAACCGAATCGATCGCACCAATTTTGAATGGAATCCCTGAT 2580
Db 2521 CCAATCGGAAGTCCGGAGAACCGAATCGATCGCACCAATTTTGAATGGAATCCCTGAT 2580
Qy 2581 CTAGATTTGTTCTCGAGAGATGGAGAAATATGTCGGCATATTTCCCATCAATTTCTTTG 2640
Db 2581 CTAGATTTGTTCTCGAGAGATGGAGAAATATGTCGGCATATTTCCCATCAATTTCTTTG 2640
Qy 2641 GATATTCATATTTGATGTCACAGACTTCATGAGATCTAGCGGTGTCGGTATTCAG 2700
Db 2641 GATATTCATATTTGATGTCACAGACTTCATGAGATCTAGCGGTGTCGGTATTCAG 2700
Qy 2701 ATTAAGACGCGAGGAGTCTATGCAAGACTAGCGAATCTGGAATTTATGAAGAGAAACCA 2760
Db 2701 ATTAAGACGCGAGGAGTCTATGCAAGACTAGCGAATCTGGAATTTATGAAGAGAAACCA 2760
Qy 2761 TTATTAGGAGAGCACTGTCTCGTGTGAAGAGACAGAGAAAAAATGGAGAGACAAACGT 2820
Db 2761 TTATTAGGAGAGCACTGTCTCGTGTGAAGAGACAGAGAAAAAATGGAGAGACAAACGT 2820
Qy 2821 GAAAACTACAAATTTGAAACAAACGATATATACAGAGGCAAAAGAGCTGTGATGCT 2880
Db 2821 GAAAACTACAAATTTGAAACAAACGATATATACAGAGGCAAAAGAGCTGTGATGCT 2880
Qy 2881 TTATTGTAGATTTCTCAATATAATAGATTACAACGGATACAAACATTTGGCATGATTCAT 2940
Db 2881 TTATTGTAGATTTCTCAATATAATAGATTACAACGGATACAAACATTTGGCATGATTCAT 2940
Qy 2941 GCGGCAGATAAATTTGTTATCGAATTTGAGAGGCTTATCTGTGAGAAATTTATCTGTTATC 3000
Db 2941 GCGGCAGATAAATTTGTTATCGAATTTGAGAGGCTTATCTGTGAGAAATTTATCTGTTATC 3000
Qy 3001 CCGGGTGAATATGCGGAAATTTTGAAGAAATAGAGGTGCGATATCATCTGCAATCTCC 3060
Db 3001 CCGGGTGAATATGCGGAAATTTTGAAGAAATAGAGGTGCGATATCATCTGCAATCTCC 3060
Qy 3061 CTATACGATGCGAGAAATGTCGTTAAATAATGTCATTTTAATGATGATGATGCTGCG 3120
Db 3061 CTATACGATGCGAGAAATGTCGTTAAATAATGTCATTTTAATGATGATGATGCTGCG 3120
```

```
Qy 3121 AATGTAAGGCGATGTAGATGTACAAAGAGCCATCACCGTTCTGCTCTGTTATCCCA 3180
Db 3121 AATGTAAGGCGATGTAGATGTACAAAGAGCCATCACCGTTCTGCTCTGTTATCCCA 3180
Qy 3181 GAATGGGAAGCAGAAAGTGTCAAGCAGTTTCGCGTCTGTCCGGGGCGTGGCTATATCCTC 3240
Db 3181 GAATGGGAAGCAGAAAGTGTCAAGCAGTTTCGCGTCTGTCCGGGGCGTGGCTATATCCTC 3240
Qy 3241 CGTGTACACGGGTACAAAGAGGGATATGGAGAGGGTTGTGTAAACGATCCATGAATCGAG 3300
Db 3241 CGTGTACACGGGTACAAAGAGGGATATGGAGAGGGTTGTGTAAACGATCCATGAATCGAG 3300
Qy 3301 AACAAATACAGCAACTTAAATTTAAAACTGTGAAGAGAGGAAGTGTATCCAAACGGAT 3360
Db 3301 AACAAATACAGCAACTTAAATTTAAAACTGTGAAGAGAGGAAGTGTATCCAAACGGAT 3360
Qy 3361 ACAGGAACGTTGTAATGATTAATCTACGACACCAAGGTACAGCAGTATGTAATTCGCTAAT 3420
Db 3361 ACAGGAACGTTGTAATGATTAATCTACGACACCAAGGTACAGCAGTATGTAATTCGCTAAT 3420
Qy 3421 GCTGGATATGAGGATGATATGAAGTTGATCTACAGCATCTGTAAATTTACAAACCGACT 3480
Db 3421 GCTGGATATGAGGATGATATGAAGTTGATCTACAGCATCTGTAAATTTACAAACCGACT 3480
Qy 3481 TATCAAGAGAAACGTTATACAGATGTACAGAGAGATATCATTTGTAATGACAGAGGG 3540
Db 3481 TATCAAGAGAAACGTTATACAGATGTACAGAGAGATATCATTTGTAATGACAGAGGG 3540
Qy 3541 TATGTGAATTTATCCACCACCTACCGCTGTTATATGACAAAGAAATTTAGAAATCTTCCCA 3600
Db 3541 TATGTGAATTTATCCACCACCTACCGCTGTTATATGACAAAGAAATTTAGAAATCTTCCCA 3600
Qy 3601 GAAACCGATAGGTATGATTTGAGATTGGAGAAACGGAAGGAAATTTATTTGTAGACAGC 3660
Db 3601 GAAACCGATAGGTATGATTTGAGATTGGAGAAACGGAAGGAAATTTATTTGTAGACAGC 3660
Qy 3661 GTGGAAATTTACTCTTATGGAGGAATAG 3687
Db 3661 GTGGAAATTTACTCTTATGGAGGAATAG 3687

RESULT 2
AAS02482
ID AAS02482 standard; DNA; 4173 BP.
XX
AC AAS02482;
XX
DT 29-AUG-2001 (first entry)
XX
DE B. thuringiensis DNA encoding a toxic crystal protein, CryET64.
XX
KW Delta endotoxin; Lepidopteran-active; crystal protein; insecticide;
KW transgenic plant; corn; wheat; soybean; oat; cotton; rice; y; sorghum;
KW sugarcane; tomato; tobacco; kapok; flax; potato; barley; turf grass;
KW pasture grass; berry; fruit; legume; vegetable; ornamental plant; shrub;
KW cactus; tree cell; gypsy moth; looper; tobacco budworm; spruce budworm;
KW cotton leaf perforator; CryET64; ds.
XX
OS Bacillus thuringiensis.
XX
FH Key Location/Qualifiers
FT CDS 1..3471
FT FT /*tag= a
FT FT /product= "CryET64"
FT FT /transl_except= (pos:1..3,aa:Met)
XX
XX WO200119859-A2.
XX
XX 22-MAR-2001.
XX
XX 13-SEP-2000; 2000WO-US025361.
XX
XX 15-SEP-1999; 99US-0153995P.
PR
```


DE Bacillus thuringiensis 158C2c toxin encoding DNA.
 KW Bacillus thuringiensis; 158C2c; toxin; lepidopteran-active; insect;
 KW microbe; transgenic plant; resistant; ss.
 XX Bacillus thuringiensis.
 OS
 FH Key Location/Qualifiers
 FT CDS 1..3684
 FT /*tag= a
 FT /product= "158C2c toxin"
 XX
 PN US5723758-A.
 XX
 XX 03-MAR-1998.
 XX
 XX 23-MAY-1995; 95US-00448170.
 XX
 XX 13-SEP-1991; 91US-00759247.
 PR 01-JUN-1993; 93US-00069902.
 XX
 XX (MYCO) MYCOGEN CORP.
 XX
 XX Steelman S, Narva KE, Cummings DA, Payne J, Cannon RJ;
 PI WPI; 1998-178600/16.
 DR P-PSDB; AAW44321.
 XX
 XX DNA encoding Bacillus thuringiensis toxin proteins - for producing
 FT transgenic plants resistant to attack by lepidopteran pests.
 XX
 XX Claim 7; Col 29-34; 25pp; English.
 XX
 CC The present sequence encodes a 158C2c toxin from Bacillus thuringiensis.
 CC The present invention describes isolated nucleic acid sequences encoding
 CC toxins active against lepidopteran insects. The present invention also
 CC describes a recombinant host transformed with a nucleic acid sequence
 CC encoding a toxin. The nucleic acid sequence can be amplified. The product
 CC is used for producing transgenic plants resistant to attack by
 CC lepidopteran pests. N.B. The numbers in the sequence listing do not
 CC correspond to the SEQ ID NO: s mentioned in the specification e.g. in the
 CC specification SEQ ID NO:3, 4, 5 and 6 are said to correspond to the
 CC nucleic acid sequences encoding the toxins, and SEQ ID NO:7, 8, 9 and 10
 CC are said to correspond to the amino acid sequences of the toxins, whereas
 CC in the sequence listing the amino acid sequence is given after the
 CC nucleic acid sequence encoding it (i.e. SEQ ID NO:3 encodes SEQ ID NO:4)
 XX
 XX Sequence 3684 BP; 1232 A; 628 C; 803 G; 1021 T; 0 U; 0 Other;
 Query Match 88.6%; Score 3265.6; DB 2; Length 3684;
 Best Local Similarity 93.4%; Pred. No. 0;
 Matches 3449; Conservative 0; Mismatches 229; Indels 15; Gaps 3;
 QY 1 TTGACTTCAAAATAGGAAAATGAGATGAAATATATAATGCTTTATCGATTCAGCTGTA 60
 DB 1 TTGACTTCAAAATAGGAAAATGAGATGAAATATATAATGCTTTATCGATTCAGCTGTA 60
 QY 61 TCGAATCATTTCCACAAATGGAATCTATCACAGATGCTCGTATTCAGGATTTCTTGTTG 120
 DB 61 TCGAATCATTTCCGCAAAATGGAATCTATCAACGATGCTCGTATTCAGGATGATCTTGTTG 120
 QY 121 ATAGCCGAGGGGAATATATCAATTCACCTTTGTAGGCGCATCAACAGTCCAAACGGGTATT 180
 DB 121 ATAGCCGAGGGGAACAAATATCGATCCATTTGTAGGCGCATCAACAGTCCAAACGGGTATT 180
 QY 181 AACATAGCTGGTGAATATACATAGTGTATTAGGCGTACCGTTTCTGCTGGAACAAATAGTAGT 240
 DB 181 AACATAGCTGGTGAATATACATAGTGTATTAGGCGTACCGTTTCTGCTGGAACAAATAGTAGT 240
 QY 241 TTTTATAGTTTCTCTGTTGTTGTAATATGCGCGGCGGAGATCAGTGGGAAATTTTC 300
 DB 241 TTTTATAGTTTCTCTGTTGTTGTAATATGCGCGGCGGAGATCCTTGGGAAATTTTC 300

QY 301 CTAGAAACATGTGCAACAACTTATAAATCAACAAATACAGAAAATGCTAGGAATACGGCA 360
 DB 301 CTAGAACATGTGCAACAACTTATAAACAACAACAGTAAACAGAAAATACCTAGGATACGGCT 360
 QY 361 CTTGCTCGAATTACAAGGTTTAGGAAATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT 420
 DB 361 CTTGCTCGAATTACAAGGTTTAGGAAATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT 420
 QY 421 TGGCTAGAAAACCGTGATGATGCAAGAACGAGAAAGTGTCTTTTATACCAATATATAGCC 480
 DB 421 TGGCTAGAAAACCGTGATGATGCAAGAACGAGAAAGTGTCTTTTATACCAATATATAGCC 480
 QY 481 TTAGAACTTGAATTTCTTAATGCGATGCGCTTTTCGCAATTAGAAAACCAAGATTCCA 540
 DB 481 TTAGAACTTGAATTTCTTAATGCGATGCGCTTTTCGCAATTAGAAAACCAAGATTCCA 540
 QY 541 TTATTAATGGTATATGCTCAAGTGCAGAAATTTACACCTATTTATTTAGAGATGCTCT 600
 DB 541 TTATTAATGGTATATGCTCAAGTGCAGAAATTTACACCTATTTATTTAGAGATGCTCT 600
 QY 601 CTTTTTGGTAGTCAATTTGGGCTTACATCGCAGGAAATTCACGGTTATTATGAGCGCCAA 660
 DB 601 CTTTTTGGTAGTCAATTTGGGCTTACATCCCAAGAAATTCACGGTTATTATGAGCGCCAA 660
 QY 661 GTGGAACAAACGAGAGATTATTCCGACTATTGCGTAGAATGGTATTAATACAGTCTAAAT 720
 DB 661 GTGGAACAAACGAGAGAAATTTCTGATTATTGCGAAGATGGTATTAATACCGGTTAAAT 720
 QY 721 AGCTTGAGAGGACAAATCCGCAAGTGGGCGGTTTAAATCAATTCGCTAGAGATCTA 780
 DB 721 AATTTGAGAGGACAAATGCTGAAAGTGGTTCGCAATATAATCAATTCGCTAGAGACTTA 780
 QY 781 AGCTTAGGGGTATTAGATCTAGTGGCACTATTTCCCAAGCTATGACACTCGCACTTATCCA 840
 DB 781 AGCTTAGGGGTATTAGATCTAGTGGCACTATTTCCCAAGCTATGACACCGGTTTATCCA 840
 QY 841 ATAAATACGAGTGTCTCAGTTTAAACAGGAAATTTATACAGACGCAATTCGAGCAACAGGG 900
 DB 841 ATGAATACCAGTGTCTCAATTTAAACAGAGAAATTTATACAGATCCCAATTCGAGCAACAAAT 900
 QY 901 GTAAT-----ATGCAAGTATGAATTCGTATATAATAATGACACTTCGTTTCGGCT 954
 DB 901 GCACCTTCAGGATTTGCAAGTACGAATTCGTTTAAATAATGACCACTCGTTTCCTGCC 960
 QY 955 ATAGAGACTCGGTTATCGAAGCCGCACTACTTTGATTTCTAGAACAACTTCAAT 1014
 DB 961 ATAGAGCTCGGTTATAGGCTCCGCACTACTTTGATTTCCAGAACAGCTTCAAT 1020
 QY 1015 TTAGCACTTCATCAGATGGAGTGTCTACTAGGCATATCACTTATGCGGGGGGACACA 1074
 DB 1021 TTGAGCGTATTAAAGTCGATGGAGTAACTCAATATATGAATTTACTGGGTGGGACATAGA 1080
 QY 1075 ATTCAATCTCGGCAATAGAGCGGATTAATACTCTCAACGATGGGTCTACCACTACT 1134
 DB 1081 CTTGAATCCGAAACAAATAAGGGGTCATTAAGTACCTCCGACACACGGAATACCAATACT 1140
 QY 1135 TCTATTAACTCTGTAAGATTATCATTTCTCTCGAGACGTATATTGGACTGAATCATAT 1194
 DB 1141 TCTATTAACTCTGTAACATTCAGTTCACTCTCGAGAGCTTTATAGACAGATCATTT 1200
 QY 1195 GCAGGAGTCTTCTATGGGAAATTTACCTTGAACCTTATTCATGGTGTCCCTACTGTAGA 1254
 DB 1201 GCAGGGATAAATAT-----ACTTCTAACTACTCTCTGTAATGGAGTACTCTGGGCTAGA 1254
 QY 1255 TTTAATTTTAGAACCTCAGAACTACTTTTGAAGAGGTACTGCTTAACTATAGTCAACCC 1314
 DB 1255 TTTAATTTGAGAAATCCCTGAAATCTCTT---AGAGGTAGCCTCTCTATATCTAGGG 1311
 QY 1315 TATGAGTCACTCGGGCTTCAATTTAAAGATTTCAGAACTGAAATTCACACAGAAAACAA 1374
 DB 1312 TATAGTGGAGTGGGACAACTATTTGATTGAGAACTGAATTCACACAGAAAACAA 1371
 QY 1375 GAACGACCAAAATATGAATCATATAGTCAATAGGTTATCTCACATAGGGGTCTATTCACAA 1434

Db 1372 GAACGACCAATATATGATCTTACAGTCATAGATATCTAATATAGACTAATATCAGGA 1431
Qy 1435 TCTAGGGTGCATGACAGTATATCTTGGACGACCCGCTAGTCAGATCGTACAAATPACC 1494
Db 1432 AACACTTTGAGAGCACCAAGTATATCTTGGACGACCCGCTAGTCAGATCGTACAAATPACC 1491
Qy 1495 ATTAGTTTCAGATAGCATTAACACAAATACCATTTGGTAAATCATTCACCTTAATTCAGGT 1554
Db 1492 ATTAGTTTCAGATAGCATTAACACAAATACCATTTGGTAAATCATTCACCTTAATTCAGGT 1551
Qy 1555 ACCTCTGTAGTCAGTGGCCACAGGATTTACAGGAGGGGATATAATCCGAACTAAAGCTTAAT 1614
Db 1552 ACCTCTGTAGTCAGTGGCCACAGGATTTACAGGAGGGGATATAATCCGAACTTAAGCTTAAT 1611
Qy 1615 GGTAGTGTACTAAGTATGGGTCTTAAATTTTAAATAATACATCATTTACAGCGGTATCGCGTG 1674
Db 1612 GGTAGTGTACTAAGTATGGGTCTTAAATTTTAAATAATACATCATTTACAGCGGTATCGCGTG 1671
Qy 1675 AGAGTTCTGTATGCTGCTTCTCAAAACAAATGGTCTGAGGGTAACTGCGGAGGGAGTACT 1734
Db 1672 AGAGTTCTGTATGCTGCTTCTCAAAACAAATGGTCTGAGGGTAACTGCGGAGGGAGTACT 1731
Qy 1735 ACTTTTCATCAAGGATTCCTGTAGTATCTAGTGCATAGTGCATAGTCTTTGACATCTCAATCA 1794
Db 1732 ACTTTTCATCAAGGATTCCTGTAGTATCTAGTGCATAGTGCATAGTCTTTGACATCTCAATCA 1791
Qy 1795 TTTAGATTTGCAGAAATTCCTGTAGTATCTAGTGCATAGTGCATAGTCTTTGACATCTCAATCA 1854
Db 1792 TTTAGATTTGCAGAAATTCCTGTAGTATCTAGTGCATAGTGCATAGTCTTTGACATCTCAATCA 1851
Qy 1855 AGTATAAGTAAATATGACAGTAGCAAAAGCTTTTCACTTTGATAAATTAAGTAAATCAATTCGA 1914
Db 1852 AGTATAAGTAAATATGACAGTAGCAAAAGCTTTTCACTTTGATAAATTAAGTAAATCAATTCGA 1911
Qy 1915 ATTACTGCAACCTTCGAAGCAGAGTAATCGATTTAGAAAGGCGCAAGAGGCGGTGAATGCT 1974
Db 1912 ATTACTGCAACCTTCGAAGCAGAGTAATCGATTTAGAAAGGCGCAAGAGGCGGTGAATGCT 1971
Qy 1975 CTGTCTTACTAATACGAATCCAGAAGATTTGAAACAGATGTGACAGATTTATCATATTTGAT 2034
Db 1972 CTGTCTTACTAATACGAATCCAGAAGCTTTGAAACAGATGTGACAGATTTATCATATTTGAT 2031
Qy 2035 CAAGTATCCAAATTTAGTGGCGTGTATTCGGATGAATCTTCTAGATGAATGAAGAAAGAGAGAA 2094
Db 2032 GAAGTATCCAAATTTAGTGGCGTGTATTCGGATGAATCTTCTGATGAATGAAGAAAGAGAGAA 2091
Qy 2095 TTACTTTGAGAAAGTGAATATGCAACGACTCAGTGCATGAAAGAACTTACTTCCAGAT 2154
Db 2092 TTACTTTGAGAAAGTGAATATGCAACGACTCAGTGCATGAAAGAACTTACTTCCAGAT 2151
Qy 2155 CCAAACTTCACATCCATCAATAGCAACACAGACTTCATATCTACTAATGAGCAATCGAAT 2214
Db 2152 CCAAACTTCACATCCATCAATAGCAACACAGACTTCATATCTACTAATGAGCAATCGAAT 2211
Qy 2215 TTCATCTATCCATGAACAACTTGAAACATGGATGGTGGGAAAGTGAAGAACTTACAAATC 2274
Db 2212 TTCATCTATCCATGAACAACTTGAAACATGGATGGTGGGAAAGTGAAGAACTTACAAATC 2271
Qy 2275 CAGGAGGAATGACGTATTTAAAGAGAAATTCGTACACTACCGGGGACTTTTAAATGAG 2334
Db 2272 CAGGAGGAATGACGTATTTAAAGAGAAATTCGTACACTACCGGGGACTTTTAAATGAG 2331
Qy 2335 TGTATTCGACGTATTTATCAAAAATAGGAGAGTCCGAAATTAAGGCTTATACTCCG 2394
Db 2332 TGTATTCGACGTATTTATCAAAAATAGGAGAGTCCGAAATTAAGGCTTATACTCCG 2391
Qy 2395 TACCAATTAAGAGGGTATATTTGAAGTAGTCAAGATTTAGAGATATATTTGATTCGTTAT 2454
Db 2392 TACCAATTAAGTGGCTATATTTGAAGTAGTCAAGATTTAGAGATATATTTGATTCGTTAT 2451
Qy 2455 AATCGGAAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTCAGTT 2514

Db 2452 AATCGGAAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGATATGCGCGCTTTCAGTT 2511
Qy 2515 GAAAGCCCAATTCGAAAGGTGCGAGAACCGAATCGATCGCACACACATTTTGAATGGAAT 2574
Db 2512 GAAAGCCCAATTCGAAAGGTGCGAGAACCGAATCGATCGCACACACATTTTGAATGGAAT 2571
Qy 2575 CCTGATGTAGATTTCTTCGACAGATGGAGAAATTTGGGATCATTTCCCATCATTTTC 2634
Db 2572 CCTGATGTAGATTTCTTCGACAGATGGAGAAATTTGGGATCATTTCCCATCATTTTC 2631
Qy 2635 TCTTTGGATATTCATATTTGGATGCACAGACTTCGATGAGAAATCTAGGCGTGTGGGTGGTA 2694
Db 2632 TCTTTGGATATTCATATTTGGATGCATAGACTTCGATGAGAACTTAGGCGTGTGGGTGGTA 2691
Qy 2695 TTCAAGATTAAGACGCGAGGAAAGTCTCAAGACTAGGAAATCTGGAATTTATTTGAAGAG 2754
Db 2692 TTCAAGATTAAGACGCGAGGAAAGTCTCAAGACTAGGAAATCTGGAATTTATTTGAAGAG 2751
Qy 2755 AAACCATTTATTAGGAGAACCACTGCTCTGCTGTAAGAGAGCAGAGAAATTTGGAGAGAC 2814
Db 2752 AAACCATTTATTAGGAGAACCACTGCTCTGCTGTAAGAGAGCAGAGAAATTTGGAGAGAC 2811
Qy 2815 AAACGTGAAAACTCAATTTGGAAACAAACAGAGTATATACAGAGGCAAAAGAGCTGTG 2874
Db 2812 AAACGTGAAAACTCAATTTGGAAACAAACAGAGTATATACAGAGGCAAAAGAGCTGTG 2871
Qy 2875 GATGCTTTTATTTGTAGATCTCAATATATAGATTTCAAGCGGATACAAACATTTGGCATG 2934
Db 2872 GATGCTTTTATTTGTAGATCTCAATATATAGATTTCAAGCGGATACAAACATTTGGCATG 2931
Qy 2935 ATTCTATCGGCGAGATAAATCTTGTTCATCGAATTTGAGAGGCTTATCTGTGAGAAATTTCT 2994
Db 2932 ATTCTATCGGCGAGATAAATCTTGTTCATCGAATTTGAGAGGCTTATCTTTCAGAAATTTCT 2991
Qy 2995 GTTATCCCGGTGTAATTCGGAATTTTGAAGAAATTTGAAGAAATTTGAAGGTCGATTTACCTGCA 3054
Db 2992 GTTATCCCGGTGTAATTCGGAATTTTGAAGAAATTTGAAGAAATTTGAAGGTCGATTTACCTGCA 3051
Qy 3055 ATCTCCCTATACGATCGGAGAAATGTCTGTAAATAATGGTGAATTTAATAATGGAATGACA 3114
Db 3052 ATCTCCCTATACGATCGGAGAAATGTCTGTAAATAATGGTGAATTTAATAATGGAATGACA 3111
Qy 3115 TGTCTGAATTTGAAAGCGGATGTAGATGTACAACAGAGCCATCACCGTTCCTGCTTGT 3174
Db 3112 TGTCTGAATTTGAAAGCGGATGTAGATGTACAACAGAGCCATCACCGTTCCTGCTTGT 3171
Qy 3175 ATCCAGAAATGGAAGCAGAGGTGTCAAGCAGTTCGCGTCTGTCGCGGCGGTGGCTAT 3234
Db 3172 ATCCAGAAATGGAAGCAGAGGTGTCAAGCAGTTCGCGTCTGTCGCGGCGGTGGCTAT 3231
Qy 3235 ATCTCCCGTGTCAACGCGTACAAAGAGGGATATGGAGAGGGTGTGTAAACGATCCCATGAA 3294
Db 3232 ATCTCCCGTGTCAACGCGTACAAAGAGGGATATGGAGAGGGTGTGTAAACGATCCCATGAA 3291
Qy 3295 ATCGAGAACATATACAGACGAACTAAATTTAAATACTGTGAAGAGAGGAGGTGTATCCA 3354
Db 3292 ATCGAGAACATATACAGACGAACTAAATTTAAATACTGTGAAGAGAGGAGGTGTATCCA 3351
Qy 3355 ACGGATACAGGAGCGTGAATGATTTACTGCACACCAAGGTACACGCGTATCTGAATTC 3414
Db 3352 ACGGATACAGGAGCGTGAATGATTTACTGCACACCAAGGTACAGCGATGTGAATTC 3411
Qy 3415 CGTAATGCTGGATATGAGGATGATATGAATTTGATATCTACAGCATCTGTAAATTTACAAA 3474
Db 3412 CGTAATGCTGGATATGAGGATGATATGAATTTGATATCTACAGCATCTGTAAATTTACAAA 3471
Qy 3475 CCGACTTTATGAAGAGAAACGATATACAGATGTACGAAGAGATTAATCATTTGTGAATATGAC 3534
Db 3472 CCGACTTTATGAAGAGAAACGATATACAGATGTACGAAGAGATTAATCATTTGTGAATATGAC 3531
Qy 3535 AGAGGCTATGTGAATTTATCCACCACTACCGCTGGTTATATGACAAAGAAATTTAGAAATAC 3594
Db 3532 AGAGGCTATGTGAATTTATCCACCACTACCGCTGGTTATATGACAAAGAAATTTAGAAATAC 3591

QY 3595 TTCCGAAACCGATAAGTATGATTGAGATTGAGAAACGGAAGGAAGTTTATTGTA 3654
 DB 3592 TTCCGAAACCGATAAGTATGATTGAGATTGAGAAACGGAAGGAAGTTTATTGTA 3651
 QY 3655 GACAGCTGGAATTAATCTCTTATGAGGAATAG 3687
 DB 3652 GACAGCTGGAATTAATCTCTTATGAGGAATAG 3684

RESULT 4

AAA89221 ID AAA89221 standard; DNA; 3684 BP.

AC AAA89221;

XX 19-MAR-2001 (first entry)

DT Bacillus thuringiensis delta-endotoxin gene 158C2c.

DE Delta-endotoxin; 158C2c; toxin; lepidoptera; crystal protein;
 KW transgenic plant; insect resistance; crop protection; biological control;
 KW ss.

XX Bacillus thuringiensis.

OS US6150589-A.

PN 21-NOV-2000.

XX 31-OCT-1997; 97US-00961803.

XX 23-MAY-1995; 95US-00448170.

XX (MYCO) MYCOGEN CORP.

XX Narva KE, Stelman S, Payne J, Cummings DA, Cannon RJC;

XX WPI; 2001-049107/06.

DR P-PSDB; AAB19950.

PT New genes isolated from Bacillus thuringiensis which encode toxins
 PT against lepidopteran insects and which can be used to transform various
 PT hosts to express the B.t. toxin and confer insect resistance.

XX Example 2; Col 19-24; 25pp; English.

XX This is the partial nucleotide sequence of the novel delta-endotoxin
 CC 1582Cc gene of Bacillus thuringiensis (B.t.) isolate PS158C2 (NRRL B-
 CC 19872). To isolate novel toxin genes from PS158C2, total cellular DNA was
 CC subjected to PCR amplification using primers (see AAA89222-23) based on
 CC conserved sequences of B.t. toxins. The PCR products were used as probes
 CC to screen a PS158C2 gene library. Plasmid subclone pMYC2383 was obtained
 CC which contained the entire coding sequence for 158C2c (see AAA89221), the
 CC 3' portion of the 158C2b gene (see AAA89220) and the 5' portion of 158C2d
 CC (see AAA89218). The 158C2a gene was isolated on a separate subclone,
 CC pMYC2387. These novel genes encode delta-endotoxins (see AAB19947-50)
 CC that are active against lepidopteran pests. Expression in transformed
 CC plants confers resistance to such pests, while transformed microorganisms
 CC expressing the delta-endotoxins can be applied to sites where they will
 CC proliferate and be ingested by the pest

XX SQ Sequence 3684 BP; 1232 A; 628 C; 803 G; 1021 T; 0 U; 0 Other;

Query Match 88.68; Score 3266.6; DB 4; Length 3684;

Best Local Similarity 93.48; Pred. No. 0;

Matches 3449; Conservative 0; Mismatches 229; Indels 15; Gaps 3;

QY 1 TTGACTTCAATAGGAAATAGAAATGAAATATATAATGCTTTATCGATTCCAGCTGTA 60

DB 1 TTGACTTCAATAGGAAATAGAAATGAAATATATAATGCTTTATCGATTCCAGCTGTA 60

QY 61 TCGAATCAITCCACAAATGGATCTATATCCAGATGCTCGTATTGAGGATTCCTTTGTGT 120

DB 61 TCGAATCAITCCGCAAAATGAATCTATCAACCGATGCTCGTATTGAGGATAGCTGTGT 120
 QY 121 ATAGCCGAGGGGATTAATATCAATCCACTGTTTGGCGCATCAACAGTCCAAACGGGTATT 180
 DB 121 ATAGCCGAGGGGAAACAATATCGATCCATTTGTTAGCGCATCAACAGTCCAAACGGGTATT 180
 QY 181 AACATAGCTGGTAGAATACTAGGTGTATTAGGGGTACCGTTTCTCTGACAAATAGCTAGT 240
 DB 181 AACATAGCTGGTAGAATACTAGGTGTATTAGGGGTACCGTTTCTCTGACAAATAGCTAGT 240
 QY 241 TTTTATAGTTTCTTCTGTTGTTGAATTTATGCCCCCGCGCAGAGATCAGTGGGAAATTTTC 300
 DB 241 TTTTATAGTTTCTTCTGTTGTTGAATTTATGCCCCCGCGCAGAGATCAGTGGGAAATTTTC 300
 QY 301 CTAGAACATGTCGAACAACTTATAAATCAACAATCAACAATCAACAATGCTAGGATACGGCA 360
 DB 301 CTAGAACATGTCGAACAACTTATAAAGCAACAATCAACAATCAACAATGCTAGGATACGGCT 360
 QY 361 CTGCTCGATTACAAGTTTATGAGATTTCTTTTAGAGCCTTATCAACAGTCACTTGAAGAT 420
 DB 361 CTGCTCGATTACAGGTTTATGAGATTTCTTTTAGAGCCTTATCAACAGTCACTTGAAGAT 420
 QY 421 TGGCTAGAAAACCGTGAATGCAAGAACGAGAAGTGTCTTTTATATCCAAATATATAGCC 480
 DB 421 TGGCTAGAAAACCGTGAATGCAAGAACGAGAAGTGTCTTTTATATCCAAATATATAGCC 480
 QY 481 TTAGAACTTGAATTTCTTAATGCGATGCGCTTTTCGCAATTAGAAACCAAGAGTTCCA 540
 DB 481 TTAGAACTTGAATTTCTTAATGCGATGCGCTTTTCGCAATTAGAAACCAAGAGTTCCA 540
 QY 541 TTATTAATGGTATATGCTCAAGCTGCAAAATTTTACACCTATTATTATGAGAGATGCTCT 600
 DB 541 TTATTAATGGTATATGCTCAAGCTGCAAAATTTTACACCTATTATTATGAGAGATGCTCT 600
 QY 601 CTTTGTGTAGTGAATTTGGGCTTACATCGAGGAAATTCAGCTTATTATGAGCGCCAA 660
 DB 601 CTTTGTGTAGTGAATTTGGGCTTACATCCCAAGAAATTCAGCTTATTATGAGCGCCAA 660
 QY 661 GTGGAAACAAACGAGAGATTTCCGACTATTCGTTAGATGGTATATATACAGTCTTAAAT 720
 DB 661 GTGGAAACAAACGAGAGATTTCTGATTTATGCGCAAGATGGTATATATACAGGTTTAAAT 720
 QY 721 AGCTTCAGAGGAGCAAAATCCGCAAGTGGGTGCTTATAATCAATTCGTTAGAGATCTA 780
 DB 721 AATTTGAGAGGGACAAATGCTGAAAGTTGTTGCGATATAATCAATTCGTTAGAGACTTA 780
 QY 781 AGCTTAGGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGACACTCTGCACTTATCCA 840
 DB 781 AGCTTAGGAGTATTAGATCTAGTGGCACTATTCCCAAGCTATGACACGGGTGTTTATCCA 840
 QY 841 ATAAATACGAGTCTCAGTTTAAACAGGGAAGTTTATACAGACGCAATTCGAGCAACAGGG 900
 DB 841 ATGAATACCAGTGTCTCAATTTAAACAGAGAAATTTATACAGATCCAAATGGAGAACAAAT 900
 QY 901 GTAAAT-----ATGCAAGTATGAATGTTGTTTAAATAATGCACTTCGTTTCCGCT 954
 DB 901 GCACCTTCAGGATTTGCAAGTACGAATGTTTAAATAATGCACTTCGTTTCCGCT 960
 QY 955 ATAGAGACTCGGGTTATCCGAGCGCGCATCTACTTGTATTTCTTAGAACAACTTACAAT 1014
 DB 961 ATAGAGGCTCGCGTTATTAGGCGCTCGCATCTACTTGTATTTTCAGAACAGCTTACAAT 1020
 QY 1015 TTTTAGCACTTCATCAGATGGAGTGTCTACTAGGCAATATGACTTACTGCGGGGGGACACA 1074
 DB 1021 TTCAGCGTATTAGTCGATGGAGTAAATCTCAATATATGAATTTAGTGGTGGGACATAGA 1080
 QY 1075 ATTCAATCTCGGCCAATAGGAGCGGATTAATAATCACTCAACGCAAGGGGTCTACCAATCT 1134
 DB 1081 CTTGAATCGGAACAAATAAGGGGCTCATTAAGTACCTCGACACACGGAATAATCACTACT 1140
 QY 1135 TCTATTAACTCTGTAGATTAATCAATTCCTCTCTCGAGACGTATATTGGACTGATCATAT 1194

Db 721 AATTTGAGGGGCAAAATGCTGAAAGTTGGTTGGGATATATCAATTCGGTAGAGACTTA 780
Qy 781 ACGTTAGGGGATTAGTCTAGTGGCACTATPCCAAAGCTATGACACTCGCACTTATCCA 840
Db 781 ACGCTAGGAGTATTAGATCTAGTGGCACTATTCCCAAGCTATGACACGGGTGTTTATCCA 840
Qy 841 ATAAATAGGAGTCTCAGTTAAAGGGAAGTTTATACAGACGCAATTTGAGCAACAGGG 900
Db 841 ATGAATACCAAGTGTCTCAATTAACAAGAGAAATTTATACAGATCCAAATTTGGGAGAACAAAT 900
Qy 901 GTAAAT-----ATGCGAAGTATCAATTTGGTATATAATATGATGACACCTCGTTTTCCGCT 954
Db 901 GCACCTTCAGGATTTGCAAGTACGAATGGTTTAAATTAATGACCAATCGTTTTCTGCC 960
Qy 955 ATAGAGACTCGGGTTATCCGAAGCCCGCATCTACTTGAATTTCTAGAAACAACTTACAAAT 1014
Db 961 ATAGAGGCTCGCGTTATAGGCTCCGCACTACTTGAATTTCCAGAAACAGCTTACAAT 1020
Qy 1015 TTTAGCACTTCATCAAGATGAGTGTACTAGGATATGATCTTACTGGCGGGGACACA 1074
Db 1021 TTCAGCGTATTAAAGTCGAGTGAATTAATCAATATATGAATTAATGAGTGGGACATAGA 1080
Qy 1075 ATTCAATCTCGGCCAATAGGAGCGGATTAATTAATCACTCAAGCATGGTCTACCAATACT 1134
Db 1081 CTTGAATCGCGAAACAATAAGGGGGCTAATTAAGTACCTGGAACACACGCGAATAACAATACT 1140
Qy 1135 TCTATTAACTCTGTAAATTAATCAATCTCTCGAGACGTATATGGAATGGAATCATAT 1194
Db 1141 TCTATTAACTCTGTAAATTAATCAATCTCTCGAGACGTATATGGAATGGAATCATAT 1200
Qy 1195 GCAGGAGTCTCTATGAGGGAATTTACTTTGAACCTATTCAATGGTGTCTTACTGTTAGA 1254
Db 1201 GCAGGGATAAATAT-----ACTTCTAACTACTCTGTGAATGGAGTACCTTGGGCTAGA 1254
Qy 1255 TTTAAATTTAGGACCTCGAATACTTTTGAAGAGGTACTGTAATATAGTCAACC 1314
Db 1255 TTTAAATTTGAGAAATCCCGTGAATTCCTTT---AGAGGTAGCCTTCTATACTATAGG 1311
Qy 1315 TATGAGTCACTGGGCTCAATTAAGAAATTCAGAACTGAATTTACCACACAGAAACAACA 1374
Db 1312 TATAGTGGAGTGGGACACAATTAATTTGAATTCAGAACTGAATTTACCACAGAAACAACA 1371
Qy 1375 GAACGACAAATATGAATCATATAGTCTATAGTTATCTCACATAGGGCTCATTTTCAAA 1434
Db 1372 GAACGACAAATATGAATCTTACAGTCTATAGATTATCTAATATAGACTAATACAGGA 1431
Qy 1435 TCTAGGTCATGTACAGTATATTTCTTGACCGCACCGTGTAGTCAGATCGTACAAATACC 1494
Db 1432 AACACTTTGAGAGCACAGTATATTTCTTGACCGCACCGTGTAGTCAGATCGTACAAATACC 1491
Qy 1495 ATTAGTTTCAGATAGCATAACACAAATACCATTGGTAAATCAATTTCAACCTTAAATTCAGGT 1554
Db 1492 ATTAGTTTCAGATAGCATAACACAAATACCATTGGTAAATCAATTTCAACCTTAAATTCAGGT 1551
Qy 1555 ACCTCTGTAGTCAAGTGGCCAGGATTTACAGGAGGGGATATAATCCGAACTAAAGCTTAAT 1614
Db 1552 ACCTCTGTAGTCAAGTGGCCAGGATTTACAGGAGGGGATATAATCCGAACTAAAGCTTAAT 1611
Qy 1615 GGTAGTGTACTAAGTATGGGTCTTAATTTTAAATAATACATCAATTTACAGCGGTATCGGCTG 1674
Db 1612 GGTAGTGTACTAAGTATGGGTCTTAATTTTAAATAATACATCAATTTACAGCGGTATCGGCTG 1671
Qy 1675 AGAGTTGGTTATGCTGCTCTCAACAAATGGTCTGAGGGTAACTGTCCGAGGGAGTACT 1734
Db 1672 AGAGTTGGTTATGCTGCTCTCAACAAATGGTCTGAGGGTAACTGTCCGAGGGAGTACT 1731
Qy 1735 ACTTTTATCAAGGATTCCTAGTACTATGAGTGCAGAAATGAGTCTTTGACATCTCAATCA 1794
Db 1732 ACTTTTATCAAGGATTCCTAGTACTATGAGTGCAGAAATGAGTCTTTGACATCTCAATCA 1791
Qy 1795 TTTAGATTTGCAGAAATTTCTGTAGGTATTTAGTGCATCTGGCAGTCAAACTGCTGGAATA 1854

Db 1792 TTTAGATTTCCAGAAATTTCTGTAGGTATTTAGTGCATCTGGCAGTCAAACTGCTGGAATA 1851
Qy 1855 AGTATAAGTAAATATGCAGGTAGACAAACGTTTCACTTTGATAAAAATTTGAATTCATTTCCA 1914
Db 1852 AGTATAAGTAAATATGCAGGTAGACAAACGTTTCACTTTGATAAAAATTTGAATTCATTTCCA 1911
Qy 1915 ATTTACTGCAACCTTCGAAGCAGAAATACGATTTTAGAAAAAGGCGCAGAGGCGGTGAATGCT 1974
Db 1912 ATTTACTGCAACCTTCGAAGCAGAAATATGATTTTAGAAAAAGGCGCAGAGGCGGTGAATGCT 1971
Qy 1975 CTGTTTACTAATACGAATCCAAAGATTTGAAACAGATGTGACAGATTTATCATATTGAT 2034
Db 1972 CTGTTTACTAATACGAATCCAAAGGTTGAAAACAGTGTGACAGATTTATCATATTGAT 2031
Qy 2035 CAAGTATCCAAATTTAGTGGCGTCTTTATCCGATGAATTTCTGTAGATGAAAAGAGAGAA 2094
Db 2032 GAAGTATCCAAATTTAGTGGCGTCTTTATCCGATGAATTTCTGTGATGAAAAGAGAGAA 2091
Qy 2095 TTACTTTGAGAAAGTGAATATGCGAAACGACTCAGTGTATGAAAGAAACTTTACTCCAAAGT 2154
Db 2092 TTACTTTGAGAAAGTGAATATGCGAAACGACTCAGTGTATGAAAGAAACTTTACTCCAAAGT 2151
Qy 2155 CCAAACTTCAATCCATCAATTAAGCAACCAAGCTTCAATTTCTAATATGAGCAATCGAAT 2214
Db 2152 CCAAACTTCAATCCATCAATTAAGCAACCAAGCTTCAATTTCTAATATGAGCAATCGAAT 2211
Qy 2215 TTTCACTCTATCATGAAACAATCTGAACATGGATGGTGGGAAAGTGAGAACATTTACAATC 2274
Db 2212 TTTCACTCTATCATGAAACAATCTGAACATGGATGGTGGGAAAGTGAGAACATTTACAATC 2271
Qy 2275 CAGAAAGGAAATCACTATTTTAAAGAGAAATTCGTCACTACCGGGGACTTTTAAATGAG 2334
Db 2272 CAGAAAGGAAATCACTATTTTAAAGAGAAATTCGTCACTACCGGGGACTTTTAAATGAG 2331
Qy 2335 TGTATTCGACGCTTTTATATCAAAAATAGGAGAGTCCGAAATTTAAAGCTTATACTCGC 2394
Db 2332 TGTATTCGACGCTTTTATATCAAAAATAGGAGAGTCCGAAATTTAAAGCTTATACTCGC 2391
Qy 2395 TACCAATTAAGAGGTATATTGAAGATAGTCAAGATTTTAGAGATATATTGTATTCGTTAT 2454
Db 2392 TACCAATTAAGTGGCTATATTGAAGATAGTCAAGATTTTAGAGATATATTGTATTCGTTAT 2451
Qy 2455 AATCGAAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTCAAGT 2514
Db 2452 AATCGAAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTCAAGT 2511
Qy 2515 GAAAGCCCAATCGAAGGTGCGGAGAACCGAATCGATCGACACACATTTTGAATGGAAT 2574
Db 2512 GAAAGCCCAATCGAAGGTGCGGAGAACCGAATCGATCGACACACATTTTGAATGGAAT 2571
Qy 2575 CTTGATCTAGATTTCTTGCAGAGATGGAGAAAAATGTGCGCATCATTTCCCATCATTTTC 2634
Db 2572 CTTGATCTAGATTTCTTGCAGAGATGGAGAAAAATGTGCGCATCATTTCCCATCATTTTC 2631
Qy 2635 TCTTTGGATATTGATATTGGATGACAGACTTTCATGAGAAATCTAGCGGTGTGGGTGGA 2694
Db 2632 TCTTTGGATATTGATATTGGATGACATAGACTTTCATGAGAACTTAGCGGTGTGGGTGGA 2691
Qy 2695 TTTCAAGATTTAAGACGAGGAAGTCTAGCAGACTAGGGAATCTGGAATTTTATTTGAAGAG 2754
Db 2692 TTTCAAGATTTAAGACGAGGAAGTCTAGCAGACTAGGGAATCTGGAATTTTATTTGAAGAG 2751
Qy 2755 AAACCAATTTAGGAGAACCTGTCTGTGTCGAAGAGAGCAGAGAAAAAATTTGGAGAGAC 2814
Db 2752 AAACCAATTTAGGAGAACCTGTCTGTGTCGAAGAGAGCAGAGAAAAAATTTGGAGAGAC 2811
Qy 2815 AAACGTGAAAAAATACTCAATTTGAAAAAATAAGTATATACAGAGGCAAAAGAGCTGTG 2874
Db 2812 AAACGTGAAAAAATACTCAATTTGAAAAAATAAGTATATACAGAGGCAAAAGAGCTGTG 2871
Qy 2875 GATGCTTTTATTTGTAGATTTCTCAATATATAATATGANTTCAAGCGGATCAAAACATTTGGCATG 2934
Db 2872 GATGCTTTTATTTGTAGATTTCTCAATATGATATTAACAAGCGGATCAAAACATTTGGCATG 2931

QY 541 TTATTAAATGCTATATGCTCAAGCTGCAAAATTTACACCTATATATTTGAGAGATGCTCT 600
DB 711 TTATTGATGATATGCTCAAGCTGCAAAATTTACACCTATATATTTGAGAGATGCTCT 770
QY 601 CTTTTGCTAGTGAATTTGGCTTTACATCGCAGGAATTCGAAGTATATATGAGCGCAA 660
DB 771 CTTTTGCTAGTGAATTTGGCTTTACATCGCAGGAATTCGAAGTATATATGAGCGCAA 830
QY 661 GTGGAAACAAACGAGAGATTTCCGACTATGCGTAGAATGGTATATATACAGTCTTAAT 720
DB 831 GTGGAAACAAACGAGAGATTTCCGACTATGCGTAGAATGGTATATATACAGTCTTAAT 890
QY 721 AGCTTGAGAGGACAAATGCGCAAGTTGGTGGTATATATCAATTCCTGAGATCTA 780
DB 891 AGCTTGAGAGGACAAATGCGCAAGTTGGTGGTATATATCAATTCCTGAGATCTA 950
QY 781 ACCTTAGGGTATTTAGATCTAGTGGCACTATTCGAAGCTATGACACTCGCACTTATCCA 840
DB 951 ACCTTAGGGTATTTAGATCTAGTGGCACTATTCGAAGCTATGACACTCGCACTTATCCA 1010
QY 841 ATAAATACGAGTGTCTAGTTTAAACAGGAAAGTTTATACAGACCAATTTGGAGCAACAGGG 900
DB 1011 ATAAATACGAGTGTCTAGTTTAAACAGGAAAGTTTATACAGACCAATTTGGAGCAACAGGG 1070
QY 901 GTAAATATGCAAGTATGAATTTGGTATTAATTAATGCACTTCCTGCTCCATAGAG 960
DB 1071 GTAAATATGCAAGTATGAATTTGGTATTAATTAATGCACTTCCTGCTCCATAGAG 1130
QY 961 ACTCGGGTATTCGAGCCGCACTACTTTGATTTCTAGAAACAACTTACAAATTTTATG 1020
DB 1131 GCTGGGCTATTCGAGCCGCACTACTTTGATTTCTAGAAACAACTTACAAATTTTATG 1190
QY 1021 ACTTCATACGATGGAGTGTCTAGTGCATATGATTTACTGGGGGGGACACAAATTC 1080
DB 1191 GCTTCATACGATGGAGTGTCTAGTGCATATGATTTACTGGGGGGGACACAAATTC 1250
QY 1081 TCTCGGCAATAGAGCGGATTAATACCTCAAGCATGGGTCTACCAATATCTTAT 1140
DB 1251 TCTCGGCAATAGAGCGGATTAATACCTCAAGCATGGGTCTACCAATATCTTAT 1310
QY 1141 AATCCTGTAGATTTATCATTTCTCTCGAGAGCTATATTGGACTGAATCATATGAGGA 1200
DB 1311 AATCCTGTAGATTTATCATTTCTCTCGAGAGCTATATTGGACTGAATCATATGAGGA 1370
QY 1201 GTGCTTCTATGGGAAATTTACCTTGAACCTATTCATGGTGTCCCTACTGTTAGATTTAAT 1260
DB 1371 GTGCTTCTATGGGAAATTTACCTTGAACCTATTCATGGTGTCCCTACTGTTAGATTTAAT 1430
QY 1261 TTTAGGAACCTCAGAAATCTTTTGAAGAGTACTGCTAACTATATAGTCAACCTATGAG 1320
DB 1431 TTTAGGAACCTCAGAAATCTTTTGAAGAGTACTGCTAACTATATAGTCAACCTATGAG 1490
QY 1321 TCACCTGGGCTTCAATTTAAAGATTCAGAACTGAATTTACCAACAGAAACAAACAGACGA 1380
DB 1491 TCACCTGGGCTTCAATTTAAAGATTCAGAACTGAATTTACCAACAGAAACAAACAGACGA 1550
QY 1381 CCAAAATATGAATCATATAGTCTATAGTGTATCTCAATAGGGCTCAATTCACAACTTAGG 1440
DB 1551 CCAAAATATGAATCATATAGTCTATAGTGTATCTCAATAGGGCTCAATTCACAACTTAGG 1610
QY 1441 GTGATGTACAGATATTTCTTGAACGCAAGTATGAGATCGTACAAATATCAATTAGT 1500
DB 1611 GTGATGTACAGATATTTCTTGAACGCAAGTATGAGATCGTACAAATATCAATTAGT 1670
QY 1501 TCAGATAGCATACAAATACCAATTTGGTAAATCATTTCAACCTTAATTCAGGTCTCT 1560
DB 1671 CCAAAATAGAAATCACCAAAATTTCCCAATGGTAAAGCATCCCACTTCCTCAAGGTACCACT 1730
QY 1561 GTAGTCAAGTGGCCAGGATTTACAGGAGGGATATATTCGCAACTAACGTTAATTTAGT 1620
DB 1731 GTTGTAGAGGACCAAGATTTACTGGTGGGATATTTCTCGAAGACGAATATCTGGTGA 1790
QY 1621 GTACTAAGTATGGTCTTAATTTTAAATATACATCAATTAACGCGGTATCCGCTGAGATT 1680

DB 1791 TTTGGACGATAAGAGTAACCTGTTAAACGACCAATTAACCAAAAGATATCTATAGGATTC 1850
QY 1681 CGTTATGCTGCTTCTCAAAACAAATGGTCTGAGGGTAACCTGCGAGGGAGTACTACTTTTT 1740
DB 1851 CGCTATGCTTCAACTGTAGATTTTGATTTCTTGATACAGTGGAGGTACTACTGTAAAT 1910
QY 1741 GATCAAGGATTCCTCTAGTACTATGAGTGCNAATGAGTCTTTTGACATCTCAATCAATTAGA 1800
DB 1911 AATTTTAGATTTCTTACGTAACAATGAACAGTGGAGACGAATAAAATACGGAATTTTGTG 1970
QY 1801 TTTGCAAGATTTCTGCTAGTATTTAGTGCATCTGGCAGTCAA--ACTGCTGGAATAGT 1857
DB 1971 AGACGCTGCTTTTACTACACCTTTTACTTTTACAAATTCAGATATAATTCGAACGTCT 2030
QY 1858 ATAAGTAATAATCCAGGTAGACAAACGTTTTCATTTGATAAAAATTTGAATTCATTTCCAAT 1917
DB 2031 AATCAAGGCTTTAGTGGAAATGGGGAAGTGTATATAGATAAATTTGAAATTTTCCAGTT 2090
QY 1918 ACTGCAACCTTCGAAGCAGATAACGATTTAGAAAGGGCGCAAGCGGTGAATGCTCTG 1977
DB 2091 ACTGCAACCTTCGAAGCAGATAATGATTTAGAAAGAGCGCAAGAGCGGTGAATGCTCTG 2150
QY 1978 TTTACTAATACGAATCCAAAGAGATTGAAACAGATGTGACAGATTATCATATTTGATCAA 2037
DB 2151 TTTACTAATACGAATCCAAAGAGATTGAAACAGATGTGACAGATTATCATATTTGATCAA 2210
QY 2038 GTATCCAAATTTAGTGGCTGTTTATCGGATGAATTTCTGCTTAGATGAAAAGAGAGAATTA 2097
DB 2211 GTATCCAAATTTAGTGGCTGTTTATCGGATGAATTTCTGCTTAGATGAAAAGAGAGAATTA 2270
QY 2098 CTTGAGAAAGTGAAATATATGCGAAACGACTCAGTGTGATGAAAGAACTTATCTCCAGATCCA 2157
DB 2271 CTTGAGAAAGTGAAATATATGCGAAACGACTCAGTGTGATGAAAGAACTTATCTCCAGATCCA 2330
QY 2158 AACTTCACATCCATCAATTAAGCAACAGACTTCTATCTACTAATGAGCAATCGAATTTTC 2217
DB 2331 AACTTCACATCCATCAATTAAGCAACAGACTTCTATCTACTAATGAGCAATCGAATTTTC 2390
QY 2218 ACATCTATCCATGAAACAACTCTGAAACATGATGGTGGGAGTGCAGAACATTTACAATCCAG 2277
DB 2391 ACATCTATCCATGAAACAACTCTGAAACATGATGGTGGGAGTGCAGAACATTTACAATCCAG 2450
QY 2278 GAAGAAATGACGTATTTAAAGAGAAATTAAGTCACTACCGGGGACTTTTAAATAGTGT 2337
DB 2451 GAAGAAATGACGTATTTAAAGAGAAATTAAGTCACTACCGGGTACTTTTAAATAGTGT 2510
QY 2338 TATCCAGCTATTTATCAAAAATAAGAGAGTCCGAATTTAAAGCTTATCTCGCTAC 2397
DB 2511 TATCCAGCTATTTATCAAAAATAAGAGAGTCCGAATTTAAAGCTTATCTCGCTAC 2570
QY 2398 CAATTAAGAGGGTATATTGAAGATAGTCAAGATTTAGAGATATATTGATTCGTATAAT 2457
DB 2571 CAATTAAGAGGGTATATTGAAGATAGTCAAGATTTAGAGATATATTGATTCGTATAAT 2630
QY 2458 GCGAAACATGAAACATTTGGATTTCCAGGTACCGAGTCCGATGGCGCTTTTCAAGTTGAA 2517
DB 2631 GCGAAACATGAAACATTTGGATTTCCAGGTACCGAGTCCGATGGCGCTTTTCAAGTTGAA 2690
QY 2518 AGCCGAATCGGAGGTCCGAGAGAACGAAATCGATCGCACCAATTTTGAATCGAATCCCT 2577
DB 2691 AGCCGAATCGGAGGTCCGAGAGAACGAAATCGATCGCACCAATTTTGAATCGAATCCCT 2750
QY 2578 GATCTAGATTTGTTCTCGAGAGATGAGAAAAATGTGCGCATCATTTCCCATCATTTCTCT 2637
DB 2751 GATCTAGATTTGTTCTCGAGAGATGAGAAAAATGTGCGCATCATTTCCCATCATTTCTCT 2810
QY 2638 TTGGATATTTGATATTTGGATGCAAGATTTGATGAGAAATCTAGGCGGTGGGTATTC 2697
DB 2811 TTGGATATTTGATATTTGGATGCAAGATTTGATGAGAAATCTAGGCGGTGGGTATTC 2870
QY 2698 AAGATTAAGACGAGGAAGGTCTATGCAAGCTAGGAATCTGGAATTTATTGAGAGAAA 2757

Db 427 ATTGCTCGATTTAGAGGCTTAGAGAGGCTATAGATCTTACCAGCGCTCTTGAAACT 486
Qy |||||
Db 421 TGGCTAGAAAACCGGTGATGATGCAAGAACGAGAAAGTGTCTTTATATACCCTAATATATAGCC 480
Qy |||||
Db 487 TGGTTAGATAACCGAAATGATGCAAGATCAAGAGCAATATCTTGAGCGCTATGTTGCT 546
Qy |||||
Db 481 TTAGAACTTGATTTCTTAATGCGATGCCGCTTTTCGCAATTAAGAAACCAAGAGATTCCA 540
Qy |||||
Db 547 TTAGAACTTGACATTTACTACTGCTATACCGCTTTTCAGAAATACGAAATGAAGAAAGTTCCA 606
Qy |||||
Db 541 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTCAGAGATGCGCTCT 600
Qy |||||
Db 607 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGACGCATCC 666
Qy |||||
Db 601 CTTTTTGGTAGTGAATTTGGGCTTTACATCGCAGGAAATTTCAACGTTATTATTGAGCGCAA 660
Qy |||||
Db 667 CTTTTTGGTAGTGAATGGGGATGGCATCTCCGATGTTACCAATATTACCAAGAACAA 726
Qy |||||
Db 661 GTGGAACAAACGAGAGATTATCCGACTATTTGCGTAGAATGGTATAATACAGGCTCTAAAT 720
Qy |||||
Db 727 ATCAGATATACAGAGGAATATCTAAACCATTTGCGGTACAAATGGTATAATACAGGGCTAAAT 786
Qy |||||
Db 721 AGCTTGAGGGGACAAATGCCGCAAGTTGGGTGCTTATATCAATTCGCTAGAGATCTA 780
Qy |||||
Db 787 AACTTAAGAGGGACAAATGCTGAAAGTTGGTTGCGGTATATCAATTCGCTAGAGACCTA 846
Qy |||||
Db 781 ACGTTAGGGGTATTAGATCTAGTGGCACTATTTCGAAGCTATGACACTCGCACCTTATCCA 840
Qy |||||
Db 847 ACGTTAGGGGTATTAGATTTAGTTAGTCCCTATTTCGAAGCTATGATATCTCGCACCTTATCCA 906
Qy |||||
Db 841 ATAATAACGATGCTCAGTTAAACAAGGGAAGTTTATACAGCGCAATTTGGAGCAACAGGG 900
Qy |||||
Db 907 ATCAATACGAGTGTCTCAGTTAAACAAGAGAAATTTATACAGATCCAAATGGGAGAACAAAT 966
Qy |||||
Db 901 GTAAT-----ATGGCAAGTATGAAATGGTATATTAATGATCACTTCTGTTCCGCT 954
Qy |||||
Db 967 GCACCTTCAGGATTTGCAAGTACGAAATGGTTTAAATTAATGATCACTTCTGTTCCGCT 1026
Qy |||||
Db 955 ATAGAGACTGGGGTTATCCGAAGCCGCACTACTGATTTTCAGAACCACTTACAAAT 1014
Qy |||||
Db 1027 ATAGAGCTGCAATTTTCAGGCTCCGCACTCTGATTTTCAGAACCACTTACAAAT 1086
Qy |||||
Db 1015 TTTAGCACTTCATCAGATGAGGTGCTACTAGGATATGATTTACTGCGCGGGGCAACA 1074
Qy |||||
Db 1087 TACAGTGCAATCAAGCGGTGGAGTAGCACTCAACATATGAATTTATGGGTGGACATAGG 1146
Qy |||||
Db 1075 ATTCAATCTCGGCCAATAGGAGCGGATTAATACCTCAAGCATGGGTCTACCA---AT 1131
Qy |||||
Db 1147 CTTAACTTCGCGCCCAATAGGAGGGACATTAATACTCAACACAAGGACTTACTTAATAAT 1206
Qy |||||
Db 1132 ACTTCTATTAACTCTGTAAGATATCATCTCTTCGAGACGCTATTTGCACTGTAATCA 1191
Qy |||||
Db 1207 ACTTCAATTAATCTGTAAACATTAAGTTTACGTTCTCGACGCTTTATAGAACAGAAATCA 1266
Qy |||||
Db 1192 TATCAGAGTGTCTTATGGGAATTTACCTGAACTTATCATGGTGTCCCTACTGTT 1251
Qy |||||
Db 1267 AATCCAGGGACAAATAT-----ACTATTACTACTCTGTGATGAGTACCTTGGGCT 1320
Qy |||||
Db 1252 AGATTTAATTTTAGGAACCCCTCAGAAATCTTTTGAAGAGGTACTGTAACTATAGTCAA 1311
Qy |||||
Db 1321 AGATTTAATTTTAAACCCCTCAGAAATTTATGAAGAGGGCGCACTACCTACAGTCAA 1380
Qy |||||
Db 1312 CCCTATGAGTCACTGGGCTTCAATTAAGATTCAGAACTGAATTTACCAACGAAACA 1371
Qy |||||
Db 1381 CCGTATCAGGGAGTTGGGATTTCAATTTATTTAGTATTTAGTATTTACCAACGAAACA 1440
Qy |||||
Db 1372 ACAGAACGACCAATATGAATCATATAGTCAATAGTATTCTCATAGGCTCATTTCA 1431
Qy |||||
Db 1441 ACAGAACGACCAATATGAATCATATAGTCAATAGTATTCTCATATAGGACTTAATCAT 1500
Qy |||||
Db 1432 CAATCTAGGGGTGCAATGATACAGTATATTTCTTGGACGCAACCGTAGTCAGATCGTACAAAT 1491
Qy |||||

Db 1501 GGAAACACTTTTGAGAGCACCGATCTATCTTTGGACGATCTGTAGTCGAGATCGTACGAAT 1560
Qy |||||
Db 1492 ACCATTAGTTTCAGATAGCATAAACAATAATACCATTTGGTAAATCATTTCAACTTTAAATTC 1551
Qy |||||
Db 1561 ACGATTGGACCAATAGAAATTTACAAATACCATTTGGTAAAGCACTGAATCTTTCATTCA 1620
Qy |||||
Db 1552 GGTACCTCTGTAGTCAGTGGCCCHAGGATTTTACAGGAGGGGATATAATCCCGAATCAACGTT 1611
Qy |||||
Db 1621 GGTGTTACTGTGTGGAGGGCCAGGATTTTACAGTGGGGATATCTCTCGTAGAACAAAT 1680
Qy |||||
Db 1612 AATGGTAGTGTACTAAGTATGGGTCTTAAATTTTAAATTAATCATCATTTACAGCGGTATCGC 1671
Qy |||||
Db 1581 ACGGTACATTTGGAGATATACGATTAATTAATGTGCCATTTATCCCAAGATATCGC 1740
Qy |||||
Db 1672 GTGAGAGTTCGTTATGCTGCTTCTCAAAACAATGCTCGAGGGTAACTGTCCGAGGGAGT 1731
Qy |||||
Db 1741 GTAAGGATTCGTTATGCTTCTACTACAGATTTTACAAATTTTTCACGAGAAATTAATGGAACC 1800
Qy |||||
Db 1732 ACTACTTTTGTGATCAGGATTTCCCTAGTACTATGAGTGCAAATGAGTCTTTGACATCTCAA 1791
Qy |||||
Db 1801 ACTGTTAATATTTGTTAATTTCTCAAGAACTATGAATAGGGGGGATAAATTTAGAAATATAGA 1860
Qy |||||
Db 1792 TCATTTAGATTTTCAGAAATTTCTGTAGGTATTTAGTGCATCTGCCAGTCAAACTGCTGGA 1851
Qy |||||
Db 1861 AGTTTGTAGAACTGAGGATTTAGTACTCTCTTTTAAATTTTAAATGCCCAAGCACATTC 1920
Qy |||||
Db 1852 ATAAGTATAAGTAAATATGCAAGGTAGACAAACGTTTCACTTTGATAAAATTTGAATTCATT 1911
Qy |||||
Db 1921 ACATGGGTCTCAGAGTTTTTCNAATCAGGAAGTTTATATAGATAGTGCATTTGTT 1980
Qy |||||
Db 1912 CCAATTACTCAACCTTCGAAGCAGATACGATTTTAGAAAGGGCGCAAGAGCGGTGAAT 1971
Qy |||||
Db 1981 CCAGCAGAGTAAACATTTTGAGGCAGATAATGATTTTACAAAGAGCACAAAGGCGGTGAAT 2040
Qy |||||
Db 1972 GCTCTGTTTACTAATACGAATCCAGAGATTTGAAACAGATGTGCAGATTTATCATATT 2031
Qy |||||
Db 2041 GCTCTGTTTACTTCTACAAATCCAAAGAAATTTGAAAACGATGTGCAGATTTATCATATT 2100
Qy |||||
Db 2032 GATCAAGTATCCAAATTTAGTGGGTGTTTATCGATGAATTTCTGTTAGATGAAAGAGA 2091
Qy |||||
Db 2101 GACCAAGTGTCCAATATGTTGGGCATGTTTATCAGATGAATTTGCTTGGATGAGAGCGA 2160
Qy |||||
Db 2092 GAATTACTTCGAAAAGTAAATATGCGAAACGACTCAGTCAATGAAAGAAACTTTACTCCAA 2151
Qy |||||
Db 2161 GAATTAATTTGAGAAAAGTAAATATGCGAAGCACTCAGTCAATGAAAGAAACTTTACTCCAA 2220
Qy |||||
Db 2152 GATCCAAACTTCACATCCATCAATAGCAACCCAGATTTTATATCTACTATAGCAATCG 2211
Qy |||||
Db 2221 GATCCAAACTTCACATTCATCAGTGGGCAATTAAGTTTCGATCCATCGATGGCAATCA 2280
Qy |||||
Db 2212 AATTTCACTCTATCCATGAACCAATCTGAACATGGATGGTGGGAGTGAGAACATTACA 2271
Qy |||||
Db 2281 AACTTCCCTCTATTAATGAGTATCTGAACATGGATGGTGGGAGTGCGAATGTTTACC 2340
Qy |||||
Db 2272 ATCCAGGAAGAAATGACGTTATTTAAAGAAATTAAGTCACTACCGGGGACTTTTAAAT 2331
Qy |||||
Db 2341 ATTCAGGAAGGAATGACGTTATTTAAAGAAATTAAGTCACTACCGGGGACTTTTAAAT 2400
Qy |||||
Db 2332 GAGTGTATCCGACGTTATTTATATCAAAAATAGGAGGATCGGAAATTAAGGCTTATACT 2391
Qy |||||
Db 2401 GAGTGTATCCAAATTTATTTATATCAAAAATAGGAGGATCAGAAATTTAAAGCTTATACG 2460
Qy |||||
Db 2392 CGCTACCAATTAAGAGGTTATTTAGAGTAGTCAAGATTTAGAGATATATTTGATTCGT 2451
Qy |||||
Db 2461 CGCTATCAATTAAGAGGTTATTTGAAGATAGTCAAGATCTAGAGATTTTAAATTCGT 2520
Qy |||||
Db 2452 TATAATTCGCAACATGAAACATTTGATTTTCCAGGTACCGAGTCCGATGCGCGCTTTCA 2511
Qy |||||
Db 2521 TACAATGCAAGCATGAAACATTTGATTTTCCAGGTACCGATTTCCCTATGCGCGCTTTCA 2580
Qy |||||
Db 2512 GTTGAAGCCCAATCGGAAGGTCCGAGAACCGAATCGATGCGCACCAATTTTGAATGG 2571
Qy |||||
Db 2581 GTTGAAGCCCAATCGGAAGGTCCGAGAACCAAAATCGATGCGCACCAATTTTGAATGG 2640
Qy |||||

```
Qy 2572 AATCCTGATCTAGATTGTTCTGTCAGAGATGGAGAAAATGTGGCAATCATTTCCCATCAT 2631
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2641 AATCCTGATCTAGATTGTTCTGTCAGAGATGGAGAAAATGTGGCAATCATTTCCCATCAT 2700
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2632 TTCTCTTTGGATATGATATTTGGATGACAGACTTGGCATGAGATCTAGGCGTGTGGGTG 2691
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2701 TTCACTTTGGATATGATATTTGGGTGACAGACTTGGCATGAGAACCTTAGGCGTGTGGGTG 2760
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2692 GTATTCAAGATTAAAGCGCAGGAGGTTCATGCACTAGGAATCTGGAATTTATTGAA 2751
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2761 GTATTCAAGATTAAAGCGCAGGAGGTTCATGCACTAGGAATCTGGAATTTATTGAA 2820
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2752 GAGAAACCAATTATTAGGAGAGCACTGTCTGTTGGAAGAGCAGAGAAAAAATGGAGA 2811
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2821 GAGAAACCAATTATTAGGAGAGCACTGTCTGTTGGAAGAGCAGAGAAAAAATGGAGA 2880
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2812 GACAAAGTGAAGAACTTACAAATTTGGAACAAAACGAGTATATACAGAGGCAAAAGAGCT 2871
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2881 GACAAAGTGAAGAACTTACAAATTTGGAACAAAACGAGTATATACAGAGGCAAAAGAGCT 2940
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2872 GTGATCTTTATTGTTAGATTCTCAATATAATAGATTACAAGCGGATACAAACATTGGC 2931
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2941 GTGATCTTTATTGTTAGATTCTCAATATAATAGATTACAAGCGGATACAAACATTGGC 3000
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2932 ATGATTCATCGGCAGATAAACTTTGTTTCATCGAATTCGAGAGGCTTATCTGTGAGAAATTA 2991
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3001 ATGATTCATCGGCAGATAAACTTTGTTTCATCGAATTCGAGAGGCTTATCTGTGAGAAATTA 3060
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2992 TCTGTTATCCGGGTGTAATATGCGGAATTTTGAAGAAATAGAGTCCGATTAATCACT 3051
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3061 CCTGTTATCCAGGTGTAATATGCGGAATTTTGAAGAAATAGAGTCCGATTAATCACT 3120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3052 GCAATCTCCCTATACGATGCGAGAAATGTCGTTAAATGATGATTTTAATATGAAATTA 3111
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3121 GCAATGCTTTATACGATGCGAGAAATGTCGTTAAATGATGATTTTAATATGAAATTA 3180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3112 GCATGCTGGAATGTAAAGGCGATGTAGATGTATACAAAGAGCCATCACCGTCTCTGCTCTT 3171
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3181 ACATGTTGGAATGTAAAGGCGATGTAGATGTATACAAAGAGCCATCATCGTCTCTGACCTT 3240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3172 GTTATCCAGAAATGGGAAGCAGAAATGTCACAGCAGTTTCGCTCTCGGGGCGTGGC 3231
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3241 GTTATCCAGAAATGGGAAGCAGAAATGTCACAGCAGTTTCGCTCTCGGGGCGTGGC 3300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3232 TATATCTCTCGTGTACAGCGTACAAAGAGGATATGAGAGGTTTGTGTAAACGATCCAT 3291
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3301 TATATCTCTCGTGTACAGCGTACAAAGAGGATATGAGAGGTTTGTGTAAACGATCCAT 3360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3292 GAAATCGAGAAACATACAGACGAACTAAAAATTTAAAAAAGTGTGAAGAGAGGAGTGTAT 3351
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3361 GAAATCGAGAAACATACAGACGAACTAAAAATTTAAAAAAGTGTGAAGAGAGGAGTGTAT 3420
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3352 CCAACGGATACAGAAAGTGTATGATTTATATCCACACCAAGGTACAGCAGATGTATTAAT 3411
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3421 CCAACGGATACAGAAAGTGTATGATTTATATCCACACCAAGGTACAGCAGATGTATTAAT 3480
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3412 TCCCGTAAATGCTGATATGAGGTGATATGAGTGTATGATCTACAGCATCTGTTAATTAAT 3471
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3481 TCCCGTAAATGCTGATATGAGGTGATATGAGTGTATGATCTACAGCATCTGTTAATTAAT 3540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3472 AAACCGACTTTATGAAGAGAAACGTATACAGATGTACGAAGAGATATCAATTTGTGAATAT 3531
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3541 AAACCGACTTTATGAAGAGAAACGTATACAGATGTACGAAGAGATATCAATTTGTGAATAT 3600
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3532 GACAGAGGTATGTGAATTTATCCACCACTACAGCTGTTATATGACAAAGAAATTAAGAA 3591
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3601 GACAGAGGTATGTGAATTTATCCACCACTACAGCTGTTATATGACAAAGAAATTAAGAA 3660
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3592 TACTTCCAGAAACCGGATATGATTTGAGATTTGAGATTCGAGAAACCGGAGGAGTATTTATT 3651
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3661 TACTTCCAGAAACAGATACAGTATGATTTGAGATTTGAGATTCGAGAAACCGGAGGAGTATTTATT 3720
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Qy 3652 GTAGACAGCGTGAATTTACTCTTATGGAGGAATAG 3687
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3721 GTAGTAGCGTGGAATTTACTCTTATGGAGGAATAG 3756
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
ID AAZ09160 standard; DNA; 3932 BP.
XX AAZ09160;
AC AAZ09160;
XX 18-OCT-1999 (first entry)
XX
DE B. thuringiensis cryET5 toxin DNA.
XX
KW Toxin; cryET4; cryET5; insecticidal; Lepidoptera; transformed plant;
KW crystal protein; insect; ds.
XX
OS Bacillus thuringiensis.
XX
FH Key Location/Qualifiers
FT CDS 67..3756
FT /tag= a
FT /product= "cryET5"
FT /note= "No ATG start codon"
XX
XX US5942658-A.
XX
XX 24-AUG-1999.
XX
XX 24-JUN-1997; 97US-00881340.
XX
XX 29-JUL-1993; 93US-00100709.
XX
XX 30-DEC-1993; 93US-00176865.
XX
XX 07-JUN-1995; 95US-00474038.
XX
XX (MONS ) MONSANTO CO.
XX
XX Gonzalez JM, Jany CS, Tan Y, Donovan WP;
XX
XX WPI; 1999-493544/41.
XX
XX P-PSDB; AAY30923.
XX
XX Transformed plant comprising insecticidal crystal proteins.
XX
XX Claim 4; Fig 2A-J; 50pp; English.
XX
XX This invention describes novel transformed plants containing Bacillus
XX thuringiensis strain EG5847 insecticidal crystal protein genes cryET4 or
XX cryET5. cryET4 and cryET5 are novel toxin genes which produce
XX insecticidal proteins with activity against a broad spectrum of insects
XX of the order Lepidoptera. This sequence encodes the cryET5 protein
XX described in the method of the invention
XX
XX SQ Sequence 3932 BP; 1347 A; 650 C; 840 G; 1095 T; 0 U; 0 Other;
XX
XX Query Match 72.9%; Score 2687.2; DB 2; Length 3932;
XX Best Local Similarity 83.5%; Pred. No. 0;
XX Matches 3088; Conservative 0; Mismatches 593; Indels 15; Gaps 3;

Qy 1 TTGACTTCAATAGGAAAAATGAGATGAATTAATGCTTTATCGATTCAGCTGTA 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 67 TTGACTTCAATAGGAAAAATGAGATGAATTAATGCTTTATCGATTCAGCTGTA 126
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 61 TCGAATCATTCACACAAATGGATCTATCACCAGATGCTCGTATTGAGGATTTCTTTGCT 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 127 TCGAATCTTCCAGGAAATGAATCTATCACCAGATGCTCGTATTGAGATAGCTTGT 186
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 121 ATAGCCGAGGGGAATTAATATCAATCCACTTTGTAGCGCATCAACAGTCCAAACGGGTATT 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 187 GTAGCCGAGGTGAACAAATATTGATTCATTGTTAGCGCATCAACAGTCCAAACGGGTATA 246
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 181 AACATAGCTGGTAGAATACTAGGTGATTATAGGCGTACCGTTTGTCTGACAAATAGCTAGT 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


Db 247 AACATAGCTGGTAGAATATTGGCGGTATTAGGTGGCGTTTGTGCAACACTAGCTAGT 306
Qy |||||
Db 241 TTTTATAGTTTTCTTTGTTGGTGAATATTGGCCCGCGCAGAGATCAGTGGGAAATTTTC 300
Qy |||||
Db 307 TTTTATAGTTTTCTTTGTTGGGAATATTGGCTAGTGGCAGAGATCCATGGGAATTTTC 366
Qy |||||
Db 301 CTAGAACATGTCGAACAACTTATAATCAACAATAACAGAAATAGCTAGGAATACGGCA 360
Qy |||||
Db 367 CTGGAACATGTAGAACAACTTATAAGAACAAACAGTAAACAGAAAATACTAGGAATACGGCT 426
Qy |||||
Db 361 CTTGCTCGATTACAGCTTTAGGAGATTCTTTAGAGCTTATCAACAGTCACTTGAAGAT 420
Qy |||||
Db 427 ATTGCTCGATTAGAAAGTCTAGGAAGGCTTATAGATCTTACCAGAGGCTCTTGGAACT 486
Qy |||||
Db 421 TGGCTAGAAAACCGTGATGATGCAAGAACAGAAAGTGTCTTTTATACCCCAATATATAGCC 480
Qy |||||
Db 487 TGGTTAGATAACCGAAATGATGCAAGATCAAGAGCATTTCTTGAGCGCTATGTTGCT 546
Qy |||||
Db 481 TTAGAACTTGATTTTTCTTTAATGCGATGCGCTTTTCGCAATTAAGAAACCAAGAAAGTTCCA 540
Qy |||||
Db 547 TTAGAACTTGCAATTACTACTGCTATACCGCTTTTCAGAAATACGAAATGAAGAAAGTTCCA 606
Qy |||||
Db 541 TTATTATGTTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGCGCTCT 600
Qy |||||
Db 607 TTATTATGTTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGAGCGCATCC 666
Qy |||||
Db 601 CTTTTCGTAGTGAATTTGGCGTTTACATCGCAGGAAATTCAAAGTTATTATGAGCGCCAA 660
Qy |||||
Db 667 CTTTTCGTAGTGAATTTGGGGAATGGCATCTCCGATGTTAACCAATATTACCAGAACAA 726
Qy |||||
Db 661 GTGGAACAAACGAGAGATTTTCGACTATTGCGTAGAATGGTATATAACAGTGTTCTAAAT 720
Qy |||||
Db 727 ATCAGATATACAGAGGAATATTCTAACCAATTGCGTCAATGGTATATAACAGGCTAAAT 786
Qy |||||
Db 721 AGCTTGAGAGGGAACAAATGCGCAAGTTGGTGGCTTATTAATCAATTCGTTAGAGATCTA 780
Qy |||||
Db 787 AACTTAAAGAGGGAACAAATGCTGAAAGTTGGTGGCGTATTAATCAATTCGTTAGAGACTA 846
Qy |||||
Db 781 ACGTTAGGGGTATTAGATCTAGTGGCACTATTTCCAAAGCTATGACACTCGCACATTATCCA 840
Qy |||||
Db 847 ACGTTAGGGGTATTAGATTTAGTAGCCCTATTTCCAAAGCTATGATATCTCGCACATTATCCA 906
Qy |||||
Db 841 ATAAATACGATGCTCAGTTAAACAGGGAAGTTTATACAGACGCAATTGGAGCAACAGGG 900
Qy |||||
Db 907 ATCAATACGAGTGTCTCAGTTAAACAGAGAAATTTATACAGATCCAAATTGGGAGAACAAAT 966
Qy |||||
Db 901 GTAAT-----ATGGCAAGTATGAATGGTATTAATTAATGACACCTTCGTTTCCGCT 954
Qy |||||
Db 967 GCACCTTCAGGATTTGCAAGTACGAAATGGTGTAAATAATAATGACCACCTCGTTTCTGCC 1026
Qy |||||
Db 955 ATAGAGACTCGGTTATCCGAAGCCCGCATCTACTGATTTTCTAGAACAACTTACAAAT 1014
Qy |||||
Db 1027 ATAGAGCTGCGCAATTTTCAGGCTCCGCACTCTCTGATTTTCAGAACAACTTACAAAT 1086
Qy |||||
Db 1015 TTTAGCACTTCATCAGATGAGTGTCTACTAGGCAATGACTTACTGGCGGGGACACACA 1074
Qy |||||
Db 1087 TACAGTGCATCAAGCGTTGGAGTAGCACTCAACATATGAATTTGGGTGGGACATAGG 1146
Qy |||||
Db 1075 ATTCAATCTCGGCCAATAGGAGGCGGAATTAATACCTCAACGCAATGGGTCTACCA---AT 1131
Qy |||||
Db 1147 CTTAACTTCGCGCCCAATAGGAGGGAATTAATAATACCTCAACACAAAGGACTTACTTAATAAT 1206
Qy |||||
Db 1132 ACTTCTATTATCTGTGAAGATTATCACTTCTCTCGAGACGPTATTTGAGCTGAATCA 1191
Qy |||||
Db 1207 ACTTCAATTAATCTGTAAACATTAACAGTTTACGTTTACGTTCTCGAGACGTTTATAGAACAGAAATCA 1266
Qy |||||
Db 1192 TATCAGAGAGTGTCTTATGGGGAATTTACCTGAACTTATTCATGGTGTCCCTACTGTT 1251
Qy |||||
Db 1267 AATGAGGGAACAATAT-----ACTATTACTACTCTCTGTAATGGATGAGTACCTTGGGCT 1320
Qy |||||
Db 1252 AGATTTAATTTTAGGAACCCCTCAGAATCTTTTGAAGAGGATCTGCTAACTATATAGTCAA 1311
Qy |||||

Db 1321 AGATTTAATTTTATAAAACCCCTCAGAAATATTTATGAAAGAGCGCCACTACCTACAGTCAA 1380
Qy |||||
Db 1312 CCCTATCAGTCACTCGTGGCTTTCAATTTAAAAGATTAGAAAATCTGAATTTACCAACCAAGAACCA 1371
Qy |||||
Db 1381 CCGTATCAGGGAGTTGGGATTTCAATTTATTTGATTTTCAAGAACTGAATTTACCAACCAAGAACCA 1440
Qy |||||
Db 1372 ACAGAACGACAAATTTATGAATCATATAGTCAATAGTTTATCTCATATAGGCTCATTTCA 1431
Qy |||||
Db 1441 ACAGAACGACAAATTTATGAATCATATAGTCAATAGTATATCTCATATAGGACTAATCATAT 1500
Qy |||||
Db 1432 CAATCTAGGCTGCAATGATACCAAGTATATTTTGGACGCGACCTAGTGCAGATCGTCAAAAT 1491
Qy |||||
Db 1501 GGAACACACTTTGAGAGCACCAGCTTATCTTGGACGCTATCTGATGCGATCGTACGAAT 1560
Qy |||||
Db 1492 ACCATTTAGTTCAGATAGCATAAACAAATACCAATTTGTTAAATCATTTCAACCTTTAAATTTCA 1551
Qy |||||
Db 1561 ACGATTTGGAACCAATAGAAATTTACAAATACCAATTTGTTAAAGCACTGAATCTTTCAATTTCA 1620
Qy |||||
Db 1552 GGTACCTCTGTAGTCACTGCGCCAGGATTTTACAGGAGGGGATATATATCCGAACTAAACGTT 1611
Qy |||||
Db 1621 GGTGTTACTGTTGTTGGAGGCGCAGGATTTTACAGGTGGGATATCTCTGCTGAGAACAAAT 1680
Qy |||||
Db 1612 AATGGTAGTGTACTAAGTATGGGTCTTTAAATTTTAAATAATACATCATTTACAGCGGTATCGC 1671
Qy |||||
Db 1681 ACGGTAATTTGGAGATATACGATTTAAATATTAATGTGCCATTTATCCAAAGATATCGC 1740
Qy |||||
Db 1672 GTGAGAGTTCGTTATGCTGCTTCTCAAAACAAATGGTCTCGAGGGTAACTGTGCGAGGGAGT 1731
Qy |||||
Db 1741 GTAAGGATTCGTTATGCTTCTACTACAGATTTTACAAATTTTTCACAGGAAATTAATTTGAAC 1800
Qy |||||
Db 1732 ACTACTTTTGAATCAAGATTTCCCTAGTACTATGAGTGCAAATGAGTCTTTGACATCTCAAT 1791
Qy |||||
Db 1801 ACTGTTAATATTGGTAAATTTCTCAAGAACTATGAAATAGGGGGGATAAATTTAGAAATATAGA 1860
Qy |||||
Db 1792 TCATTTAGATTTTCAGAAATTTCTGTAGTATTTAGTGCATCTGSCAGTCAAACTGCTGGA 1851
Qy |||||
Db 1861 AGTTTTAGAACTGCGAGATTTTAGTACTCTCTTTTAAATTTTAAATGCCAAAGCACATTC 1920
Qy |||||
Db 1852 ATAAGTATAAGTAAATATGCAAGGTAGACAAACGTTTCACTTTGATAAAATTTGAATTTCAAT 1911
Qy |||||
Db 1921 ACATGGGTGCTCAGAGTTTTCAAATCAGGAAGTTTATATAGATAGAGTCGAATTTGTT 1980
Qy |||||
Db 1912 CCAATTTACTGCAACCTTTCGAAGCAGAAATACGATTTAGAAAAGGCGCAAGAGCGGTGAAT 1971
Qy |||||
Db 1981 CCAGCAGAGTAAACATTTGAGGCGAGATATGATTTAGAAAAGCAGCAAAAAGCGGTGAAT 2040
Qy |||||
Db 1972 GCTCTGTTTACTTAATACGAATCCAAAGAGATTTGAAAACAGATGTCGACAGATTTATCATATT 2031
Qy |||||
Db 2041 GCTCTGTTTACTTCTACAAATCCAAAGAGATTTGAAAACAGATGTCGACAGATTTATCATATT 2100
Qy |||||
Db 2032 GATCAAGTATCCAAATTTAGTGGCGTGTATTATCGGATGAATTTCTGTTAGATGAAAAGAGA 2091
Qy |||||
Db 2101 GACCAAGTGTCCAAATATGGTGGCATGTTTATCAGATGAATTTTCTTGGATGAGAGCGA 2160
Qy |||||
Db 2092 GAATTTACTTGAGAAAAGTGAATATGCGAAACGACTCAGTGAATGAAAGAACTTACTCCAA 2151
Qy |||||
Db 2161 GAATTTATTTTGAGAAAAGTGAATATGCGAAGCGACTCAGTGAATGAAAGAAACTTACTCCAA 2220
Qy |||||
Db 2152 GATCCAACTTCATCATCCATCAATTAAGCAACCGACTTTCATATCTACTTAATGAGCAATTCG 2211
Qy |||||
Db 2221 GATCCAACTTCATCATCCATCAATTAAGTTCGATCCATTCGATGAGCAATTCATCA 2280
Qy |||||
Db 2212 AATTTACATCTATCCATGAACCAATCTGAACATGATGTTGGGGGAGTGAAGAACATTTACA 2271
Qy |||||
Db 2281 AACTTCCCTCTATTAATAGCTATCTGAACATGATGTTGGGGAGTGGCAATTTTACC 2340
Qy |||||
Db 2272 ATCCAGGAAGAAATGACGTTATTTAAAGAAATTAACCTCACTACCCGGGAGCTTTTAAAT 2331
Qy |||||
Db 2341 ATTCCAGGAAGGGAATGACGTTATTTAAAGAAATTAACCTCACTACCCGGGACTTTTAAAT 2400
Qy |||||
Db 2332 GAGTGTATTCGAGCTATTTTATATCAAAAATATAGGAGAGTTCGGAATTTAAAGCTTATACT 2391
Qy |||||
Db 2401 GAGTGTATTCCAAATTTATTTATATCAAAAATATAGGAGAGTTCGGAATTTAAAGCTTATACG 2460
Qy |||||

2392 CGTACCAATTAAGAGGGTATATTGAAGATAGTCAAGATTAGAGATATATTGATTCGT 2451
 Db |||||
 2461 CGCTATCAATTAAGAGGGTATATTGAAGATAGTCAAGATTAGAGATATATTGATTCGT 2520
 Qy |||||
 2452 TATAATCGGAAACATGAAACATTGGATGTTCCAGGTACCGAGTCCGTATGCGCGCTTCA 2511
 Db |||||
 2521 TACAATGCAAGCATGAACATTGGATGTTCCAGGTACCGAGTCCGTATGCGCGCTTCA 2580
 Qy |||||
 2512 GTTGAAGCCCAATCGGAAGTGGGAGAACCGAATCGATCGCGCACACATTTTGAATGG 2571
 Db |||||
 2581 GTTGAAGCCCAATCGGAAGTGGGAGAACCGAATCGATCGCGCACACATTTTGAATGG 2640
 Qy |||||
 2572 AATCCTGATCTAGATTGTTCTCCAGAGATGGAGAAAAATGTGCGCATCAATCCCATCAT 2631
 Db |||||
 2641 AATCCTGATCTAGATTGTTCTCCAGAGATGGAGAAAAATGTGCGCATCAATCCCATCAT 2700
 Qy |||||
 2632 TTCTCTTTGATATTTGATTTGATGACACAGACTTGCATGAGAACTAGCGGTGTTGGTG 2691
 Db |||||
 2701 TTCACTTTGGATATTTGATTTGTTGGTGCACAGACTTGCATGAGAACTAGCGGTGTTGGTG 2760
 Qy |||||
 2692 GTATTCAGGATTAAGACGAGGAGGTTCATGCAAGACTAGGAACTTGGAAATTTATTGAA 2751
 Db |||||
 2761 GTATTCAGGATTAAGACGAGGAGGTTCATGCAAGACTAGGAACTTGGAAATTTATTGAA 2820
 Qy |||||
 2752 GAGAAACCATTTATTAGGAGAGCACTGCTCTGTGTGAAGAGAGCAGAGAAAAAATGGAGA 2811
 Db |||||
 2821 GAGAAACCATTTATTGAGAGAGCACTGCTCTGTGTGAAGAGAGCAGAGAAAAAATGGAGA 2880
 Qy |||||
 2812 GACAAAGTGAAGAACTACAAATTTGGAACAAACAGAGATATACAGAGGCAAAAGAGCT 2871
 Db |||||
 2881 GACAAAGTGAAGAACTACAAATTTGGAACAAACAGAGATATACAGAGGCAAAAGAGCT 2940
 Qy |||||
 2872 GTGAGTCTTTATTGTTAGATTTCTCAATATAAGATTACAAGCGGATACAACATTTGGC 2931
 Db |||||
 2941 GTGAGTCTTTATTGTTAGATTTCTCAATATAAGATTACAAGCGGATACAACATTTGGC 3000
 Qy |||||
 2932 ATGATTCATCGCGCAGATAAACTTTGTTCAATCGAATTCGAGAGGCTTATCTGTGAGAAATTA 2991
 Db |||||
 3001 ATGATTCATCGCGCAGATAAACTTTGTTCAATCGAATTCGAGAGGCTTATCTTTCAGAAATTA 3060
 Qy |||||
 2992 TCTGTTATCCGGGTGTAATGCGGAAATTTTGAAGAAATAGAGGTCCGATATATCACT 3051
 Db |||||
 3061 CCGTGTATCCCGAGTGTAAATGCGGAAATTTTGAAGAAATAGAGGTCACTATATCACT 3120
 Qy |||||
 3052 GCAATCTCCCTATACGATCGGAGAAATGTCGTTAAATAATGTTGATTTAATATGGAATTA 3111
 Db |||||
 3121 GCAATGTCCTTATACGATCGGAAATGTCGTTAAATAATGTTGATTTAATATGGAATTA 3180
 Qy |||||
 3112 GCATGCTGGAATGTAAAGGGCATGTAGATGTACAAACAGAGCCATCACCGTTCTGTCTTT 3171
 Db |||||
 3181 ACAATGTTGGAATGTAAAGGGCATGTAGATGTACAAACAGAGCCATCATCTTCTGACCTT 3240
 Qy |||||
 3172 GTTATCCAGAAATGGGAAGCAGAAATGTCACAGAGTTCGCTGTCTGCGGGCGGTGGC 3231
 Db |||||
 3241 GTTATCCAGAAATGGGAAGCAGAAATGTCACAGAGTTCGCTGTCTGCGGGCGGTGGC 3300
 Qy |||||
 3232 TATATCTCTCGTGTACAGAGTCAAGAGGGATATGAGAGGCTTGTGTAAACGATCCAT 3291
 Db |||||
 3301 TATATCTCTCGTGTACAGAGTCAAGAGGGATATGAGAGGCTTGTGTAAACGATCCAT 3360
 Qy |||||
 3292 GAAATCCAGAAACATACAGAGCACTAAATTTTAAATAATGTTGAGAGGAGGAGGTGAT 3351
 Db |||||
 3361 GAAATCCAGAAACATACAGAGCACTAAATTTTAAATAATGTTGAGAGGAGGAGGTGAT 3420
 Qy |||||
 3352 CCAACGAGATACAGAACTGATTAATATATGTCACACCAAGGTACAGCAGATGTAAT 3411
 Db |||||
 3421 CCAACGAGATACAGAACTGATTAATATATGTCACACCAAGGTACAGCAGATGTAAT 3480
 Qy |||||
 3412 TCCCGTAATCTGATATGAGATGATATGATGATGATGATGATGATGATGATGATGATGAT 3471
 Db |||||
 3481 TCCCGTAATCTGATATGAGATGATATGATGATGATGATGATGATGATGATGATGATGAT 3540

3472 AAACCCACTTATGAAGAAGAAACGTTATACAGATGTACAGATGTACAGATGTATCTGTAATAT 3531
 Db |||||
 3541 AAACCCACTTATGAAGAAGAAACGTTATACAGATGTACAGATGTATCTGTAATAT 3600
 Qy |||||
 3532 GACAGAGGGTATGTGAATTTATCCACCACCTACACAGTGGTTATATGACAAAAGAAATTAGAA 3591
 Db |||||
 3601 GACAGAGGGTATGTGAATTTATCCACCACCTACACAGTGGTTATATGACAAAAGAAATTAGAA 3660
 Qy |||||
 3592 TACTTCCCGAAGAACCGATAGGTATGAGATTGAGATTGGAGAAAACGGAAGGAAAGTTTATT 3651
 Db |||||
 3661 TACTTCCCGAAGAACCGATAGGTATGAGATTGAGATTGGAGAAAACGGAAGGAAAGTTTATT 3720
 Qy |||||
 3652 GTAGACAGCGTGAATTAATCTCTTATGAGGAATAG 3687
 Db |||||
 3721 GTAGATAGCGTGAATTAATCTCTTATGAGGAATAG 3756
 Qy |||||
 3721 GTAGATAGCGTGAATTAATCTCTTATGAGGAATAG 3756
 Db |||||
 RESULT 9
 AAT95051
 ID AAT95051 standard; DNA; 3934 BP.
 XX
 AC AAT95051;
 XX
 DT 17-FEB-1998 (first entry)
 XX
 DE DNA encoding Bacillus thuringiensis crystal toxin CryET5.
 XX
 KW EG7283; crystal toxin; CryET5; lepidopteran pest; Lymantria dispar;
 KW Ostrinia nubilalis; Pseudoplusia includens; Plutella xylostella;
 KW Spodoptera exigua; Spodoptera frugiperda; Trichoplusia ni; ds.
 XX
 OS Bacillus thuringiensis.
 XX
 FH Location/Qualifiers
 Key 67..3756
 CDS /*tag= a
 FT /product= "CryET5"
 FT
 XX
 XX US5679343-A.
 XX
 XX 21-OCT-1997.
 PD
 XX 07-JUN-1995; 95US-00474038.
 PF
 XX 29-JUL-1993; 93US-00100709.
 PR
 XX 30-DEC-1993; 93US-00176865.
 XX
 PA (MONS) MONSANTO CO.
 XX
 PI Jany CS, Gonzalez JM, Donovan WP, Tan Y;
 XX
 XX WPI; 1997-525682/48.
 DR
 XX P-PSDB; AAW35259.
 XX
 XX Lepidopteran toxic Bacillus thuringiensis crystal protein - useful to
 PT control Lepidopteran pests.
 PS
 XX Example 2; Fig 2; 50pp; English.
 CC
 XX The present sequence encodes the Bacillus thuringiensis EG7283 crystal
 CC toxin CryET5, which, optionally in association with B. thuringiensis
 CC EG7283, can be used against lepidopteran pests. CryET5 is especially
 CC useful for controlling Lymantria dispar, Ostrinia nubilalis, Pseudoplusia
 CC includens, Plutella xylostella, Spodoptera exigua, Spodoptera frugiperda
 CC and Trichoplusia ni
 XX
 XX Sequence 3934 BP; 1347 A; 650 C; 841 G; 1096 T; 0 U; 0 Other;
 SQ
 Query Match 72.9%; Score 2687.2; DB 2; Length 3934;
 Best Local Similarity 83.5%; Pred. No. 0;
 Matches 3088; Conservative 0; Mismatches 593; Indels 15; Gaps 3;
 Qy 1 TTGACTTCAATAGGAAAAATGAGATGAATTAATGCTTTATCGATTCGACTGTA 60

Db 67 TGGCTTCAATAGGAAATGAGATGAATTAATAATGCTTTATCGATTCCACGGTA 126
Qy 61 TCGAATCATTTCCACAAAATGGATCTATACCCAGATGCTCGTATGAGGATCTTTGTGT 120
Db 127 TCGAATCTTTCCAGCGAAATGAATCTATACCCAGATGCTCGTATGAGATAGCTGTGT 186
Qy 121 ATAGCCGAGGGGAATATATCAATCCACTTGTGTAGCGCATCAACAGTCCAAACGGGTATT 180
Db 187 GTAGCCGAGGTGAACAAATATGTATCCATTTGTGTAGCGCATCAACAGTCCAAACGGGTATA 246
Qy 181 AACATAGCTGTAGAACTACTAGGTGTATTAGGCGTACCGTTTGTGACAAAATAGCTAGT 240
Db 247 AACATAGCTGTAGAAATTTGGCGGTATTAGGTGTGCGGTTTGTGACAACTAGCTAGT 306
Qy 241 TTTTATAGTTTCTTGTGTGTGAATTTATGCGCGCGCGAGAGATCAGTGGGAAATTTTC 300
Db 307 TTTTATAGTTTCTTGTGTGGGAATTTATGCGCTAGTGGCAGAGATCCATGGGAAATTTTC 366
Qy 301 CTAGAACATGTCGAACTTATTAATCAACAAATAACAGAAAATGCTAGGAATACGGCA 360
Db 367 CTGGAACATGTAGAACAACTTATAAGACAAACAAAGTAAACAGAAAATCTAGGAATACGGCT 426
Qy 361 CTGCTCGATTCAAGGTTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT 420
Db 427 ATTGCTCGATTAGAGGCTTAGGAAGGCTATAGATCTTACCAGCAGGCTCTTGAACCT 486
Qy 421 TGGCTGAAAACCGTGTATGATGCAAGAAACGAGAAAGTGTCTTTATATACCAATATATAGCC 480
Db 487 TGGTTAGATTAACGAATGATGCAAGATCAAGAAAGCATTTATCTTGAGCGCTATGTGCT 546
Qy 481 TTAGAACTTGATTTTCTTAATGCGATGCCGCTTTTTCGAATTAAGAAACCAAGAGTTCCA 540
Db 547 TTAGAACTTGACATTACTACTGTCTATACCGCTTTTCAGAAATACGAAATGAAGAAGTTCCA 606
Qy 541 TTATTAATGCTATGCTCAAGCTGCAAAATTTACACCTATATATTTAGAGAGTGCCTCT 600
Db 607 TTATTAATGCTATGCTCAAGCTGCAAAATTTTACACCTATATTTATTTAGAGAGCGCATCC 666
Qy 601 CTTTTGTGTAGTGAATTTGGGCTTACATCGCAGGAATTCGAAGTTTATTTATGAGCGCAA 660
Db 667 CTTTTGTGTAGTGAATTTGGGGAATGGAATCTTCGAATTTAACCAATATTAACCAAGACAA 726
Qy 661 GTGGAACAAACGAGAGATTAATCCGACTATGCGTAGAATGGTATATAACAGTCTAAAT 720
Db 727 ATCAGATATACAGAGGAATTTCTAACCAATGCGTACAAATGGTATAATAACAGGCTAAAT 786
Qy 721 AGCTTGAGAGGACAAATGCGCAAGTTGGTGGGTATTAATCAATTCGCTAGAGATCTA 780
Db 787 AACTTAAGAGGGACAAATGCTGAAGTTGGTGGGTATTAATCAATTCGCTAGAGACCTA 846
Qy 781 ACGTTAGGGCTATTAGATCTAGTGGCACTATTTCCAGCTATGACACTGCGCACTTATCCA 840
Db 847 ACGTTAGGGCTATTAGATTTAGTAGCCCTATTTCCAGCTATGATATCTGCACTTATCCA 906
Qy 841 ATAAATACGAGTGTCTCAGTTTAAACAGGGAAGTTTATACAGACGCAATTTGAGCAACAGGG 900
Db 907 ATCAATACGAGTGTCTCAGTTTAAACAGAGAAATTTATACAGATCCAATTTGGGAGAACAAAT 966
Qy 901 GTAAAT-----ATGGCAAGTATGAATTTGGTATTAATTAATGACCTTCGTTTTCGGCT 954
Db 967 GCACCTTCAGGAATTTGCAAGTAGCAATTTGGTTTAAATAATGACCATCGTTTTCGTGCC 1026
Qy 955 ATAGAGACTCGGGTTATCCGAGCCGCACTACTTGTATTTCTAGAACCACTTACAAAT 1014
Db 1027 ATAGAGGCTGCCATTTTCAGGCGCTCCGCACTACTTGAATTTCCAGAAACAACTTACAAAT 1086
Qy 1015 TTTAGCACTTCATCACGATGGAGTGTCTACTAGGCATATGACTTACTGCGCGGGGCAACA 1074
Db 1087 TACAGTGCAATCAAGCGGTTGGATAGCACTCAACATATGAATTTATTTGGGTGGACATAGG 1146
Qy 1075 ATTCAATCTCGGCCAAATAGAGGGCGGATTAATACTCAACGATGGGTCTACCA-----AT 1131

Db 1147 CTTAACTTCGCGCCAATAGGAGGACATTAATACTCAACACAGGACTTACTTAATAAT 1206
Qy 1132 ACTTCTATTAACTCTGTAGAAATATATCTTCTCTCGAGACGTATATTTGGACTGAATCA 1191
Db 1207 ACTTCAATTAATCTCTGTAACTATACAGTTTACGCTCGAGACGCTTTATAGAACAGATCA 1266
Qy 1192 TATGACGAGGTGCTTCTATGGGAAATTTACCTTTGAACCTTATTTATGCTGCTTACTGTT 1251
Db 1267 AATGCGAGGGAACAATAT-----ACTATTTACTCTCTGTGAATGGAGTACCTTTGGGCT 1320
Qy 1252 AGATTTAAATTTTAGGAAACCTCTCAGAACTATTTTGAAGAGGTACTGCTCACTACTATAGTCAA 1311
Db 1321 AGATTTAAATTTTAAACCTCTCAGAAATTTTATGAAGAGGCGCCACTACTACTACAGTCAA 1380
Qy 1312 CCCTATAGTCACTCTGGGCTTCAATTTAAAGATTTCAAGAACTGAAATTTACCAACAGAAACA 1371
Db 1381 CCGTATCAGGGAGTTGGGATTTCAATTTATTTGATTCAGAAACTGAAATTTACCACAGAAACA 1440
Qy 1372 ACAGAAAGCAAAATTTATGNAATCATATAGTATCATAGTTATCTCATATAGGACTTAATCAT 1431
Db 1441 ACAGAAAGCAAAATTTATGAATCATATAGTATCATAGTTATCTCATATAGGACTTAATCAT 1500
Qy 1432 CAATCTAGGCTGCACTGTACCAGTATATTTCTTGACGCGCACCGTAGTCAGATCGTACAAAT 1491
Db 1501 GGAACACATTTGAGAGCAACAGTCTATTTCTTGACGCGCATCGTAGTCAGATCGTACGAT 1560
Qy 1492 ACCATTTAGTTTCAAGTAGCATTAACAAATAACCAATTTGGTAAATCATTTCAACCTTAAATCA 1551
Db 1561 ACGATTTGGACCAATAGAAATTTACAAATACCAATTTGGTAAAGCACTGAAATCTTCATTCA 1620
Qy 1552 GGTACCTCTGTAGTCTAGTGGCCAGGATTTTACAGGAGGGGATATTAATCCGAACCTAACGTT 1611
Db 1621 GGTGTTACTGTTTGTGGAGGGCCAGGATTTTACAGTGGGGATATCTCTCTAGAACAAAT 1680
Qy 1612 AATGCTAGTCTACTAGTATGGCTCTTAATTTTAAATAATACATCATTTACAGCGGTATCGC 1671
Db 1681 ACGGGTACATTTGGAGATTAACGATTAATTAATTAATTTGTCATTTATCCCAAGATATCGC 1740
Qy 1672 GTGAGAGTTGCTTATGCTCTCTCAAACTGGTCTCTGAGGGTAACTGTCGAGGGGAGT 1731
Db 1741 GTAAGGATTCGTTATGCTCTACTACAGATTTTACAATTTTTCACAGAAATTAATGGAACC 1800
Qy 1732 ACTACTTTTGTATCAAGATTTCCCTAGTACTATGAGTGCAAATGAGTCTTTTGAATCTCAA 1791
Db 1801 ACTGTTAAATTTGTTAAATTTCTCAAGAACTATGAATAGGGGGATTAATTTAGAAATATAGA 1860
Qy 1792 TCATTTTAGATTTGCAAAATTTCTGTAGTATTTAGTGCATCTGCGAGTCAAACTGCTGGA 1851
Db 1861 AGTTTTAGAACTGCGAGATTTAGTACTCTCTTTTAAATTTTAAATGCCCCAAGCAATTC 1920
Qy 1852 ATAAGTATAAGTAATTAATGCGGTAGACAAACGTTTCACTTTGATATAAATTTGAATTCATT 1911
Db 1921 ACATTTGGTGTCTCAGAGTTTTCAAATCAGGAATTTATATAGATAGAGTCTGAATTTGTT 1980
Qy 1912 CCAATTTCTGCAACCTTCGAAGCAGAAATACGATTTTGAAGGGCGCAAGAGGGCGTGAAT 1971
Db 1981 CCAGCAGAGTAACTATTTGAGGCGAATATGATTTTGAAGAGACACAAAAGGGCGGTGAAT 2040
Qy 1972 GCTCTCTTTACTTAATACGAATCCAAAGATTTCAAAACAGATGTCAGATTTATCATATT 2031
Db 2041 GCTCTCTTTACTTCTTCAAAATCCAAAGATTTCAAAACAGATGTCAGATTTATCATATT 2100
Qy 2032 GATCAAGTATCCAAATTTAGTGGGCTGTTTATCGATGAATTTCTGCTTAGATGAAGAGAGA 2091
Db 2101 GACCAAGTGTCCAAATATGTTGGCATGTTTATCAGATGAATTTTGTGATGAGAGCGA 2160
Qy 2092 GAATTTACTTTGAGAAAGTGAATAATGCGAAACGACTCAGTGTATGAAAGAACTTACTCCAA 2151
Db 2161 GAATTTATTTGAGAAAGTGAATAATGCGAAGCACTCAGTGTATGAAAGAACTTACTCCAA 2220
Qy 2152 GATCCAAACTTTCATCTCCATCAATAAGCAACGACATTTCAATCTACTATAGGCAATCG 2211
Db 2221 GATCCAAACTTTCATCTCAGTGGGCAATTAAGTTTTCGCATCCATCGATGGACAACTA 2280

	Qy	3292	GAAATCGAAGCAAAATACAGACGAACTTAAAATTTAAAACTGTGAAGAGGAGGAGTGAT	3335
	Dd	3361	GNAATCGAGAACAATACAGACGAACTAAAATTTAAAAACTGTGAAGAGGAGGAGTGAT	3420
	Qy	3352	CCAACGGATACAGGAACGTCGTAAATGATTATATCTGCACACCAGGTACAGCAGTATGTAAT	3411
	Dd	3421	CCAACGGATACAGGAACGTCGTAAATGATTATCTGCACACCAGGTACAGCAGCATGTAAT	3480
	Qy	3412	TCCCGTAATGCTGGATATGAGGATGCAATATGAAGTTGTGATACTACAGCATCTGTAAATTAC	3471
	Dd	3481	TCCCGTAATGCTGGATATGAGGATGCAATATGAAGTTGTGATACTACAGCATCTGTAAATTAC	3540
	Qy	3472	AAACCGCACTTATCAAGAGAAACGTTATACAGATGTACGAGAGATATCATCTGTGAATAT	3531
	Dd	3541	AAACCGCACTTATCAAGAGAAACGTTATACAGATGTACGAGAGATATCATCTGTGAATAT	3600
	Qy	3532	GACAGAGGGTATGTGAATTATCCACCACCTACAGCTCCAGCTGGTTATATGACAAAAGAAATTAGAA	3591
	Dd	3601	GACAGAGGGTATGTGAATTATCCACCACCTACAGCTCCAGCTGGTTATATGACAAAAGAAATTAGAA	3660
	Qy	3592	TACTTCCCAGAAACCGATAAGGTATCGGATTGAGATTGGAGAAACGGAAGGAGGTTTATT	3651
	Dd	3661	TACTTCCCAGAAACCGATAAGGTATCGGATTGAGATTGGAGAAACGGAAGGAGGTTTATT	3720
	Qy	3652	GTAGACAGCGTGGAAATTACTTCCTTATCGGAGGAATAG	3687
	Dd	3721	GTAGATAGCGTGGAACTACTTCCTCATGGAAGATAG	3756
		RESULT 10		
		AAT68434	ID AAT68434 standard; DNA; 3934 BP.	
	XX	AC	AAT68434;	
	XX	DT	25-MAR-2003 (revised)	
	XX	DT	07-JUL-1997 (first entry)	
	XX	CryET5 gene.		
	XX	CryET5; cryET4; Bacillus thuringiensis; insecticidal crystal protein; ICP; toxin; CryI protein; lepidopteran insect; insecticide; ds.		
	OS	Bacillus thuringiensis.		
	Key	Location/Qualifiers		
	CDS	67..3756		
	FT	/tag= a		
	FT	/product= "cryET5"		
	US5616319-A.			
	01-APR-1997.			
	30-DEC-1993; 93US-00176865.			
	29-JUL-1993; 93US-00100709.			
	(MONS) MONSANTO CO.			
	Gonzalez JM, Donovan WP, Tan Y, Jany CS;			
	WPI; 1997-212077/19.			
	P-PSDB; AAU17699.			
	Bacillus thuringiensis cryET5 gene encoding insecticidal protein - useful for control of lepidopteran pests.			
	Claim 2; Fig 2; 50pp; English.			
	This sequence represents the cryET5 gene of Bacillus thuringiensis (B.t.) isolate BG5847. B.t. produces inclusions during sporulation which include insecticidal crystal proteins (ICP). ICP toxins are active in insects CC			

CC only after ingestion. Once ingested, the toxic components disrupt the
 CC midgut cells, resulting in cessation of feeding, and eventually death.
 CC The CryI proteins produced by B.t. are active against lepidopteran
 CC insects. The protein encoded by this sequence, and the CryET4 protein
 CC (see AAW1700) belong to the CryI family of ICPs. This sequence can be
 CC used to transform bacteria, which are useful as insecticides against a
 CC wide range of lepidopteran pests, and can be applied to crops, soil and
 CC seeds. The encoded protein, or especially its toxic N terminal region,
 CC can be expressed in plants, to provide protection against lepidopteran
 CC pests. This sequence, or its fragments, can also be used to isolate other
 CC similar genes. (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 3934 BP; 1347 A; 650 C; 841 G; 1096 T; 0 U; 0 Other;

Query Match 72.9%; Score 2687.2; DB 2; Length 3934;
 Best Local Similarity 83.5%; Pred. No. 0;
 Matches 3088; Conservative 0; Mismatches 593; Indels 15; Gaps 3;

QY	1	TTGACTTCAAAATAGGAAATAGAAATGAATTAATTAATGCTTTATCGATTCGAGCTGTA	60
DB	67	TTGACTTCAAAATAGGAAATAGAAATGAATTAATTAATGCTTTATCGATTCGAGCTGTA	126
QY	61	TCGAATCATTTCCACAAATAGGATCTATCACCAGATGCTGTATTGAGGATCTTTTGTGT	120
DB	127	TCGAATCTTCCAGGCAATGAATCTATCACCAGATGCTGTATTGAGGATAGCTTGTGT	186
QY	121	ATAGCCGAGGGGAATAATATCAATCCACTTGTATTAGCGCATCAACAGATCCAAACGGGTATT	180
DB	187	GTAGCCGAGGTGAACAATATTGATCCATTTGTTAGCGCATCAACAGATCCAAACGGGTATA	246
QY	181	AACATAGCTGTAGAACTAGGTGTTATAGGCGTACCGTTTGTGTGACAAATAGCTAGT	240
DB	247	AACATAGCTGTAGAAATATTGGGCGTATTAGGTGTGCCGTTTGTGTGACAACTAGCTAGT	306
QY	241	TTTTATAGTTTCTTGTGTGTAATATTGCCCGCGCGAGAGATCAGTGGGAAATTTTC	300
DB	307	TTTTATAGTTTCTTGTGTGGGAATATTGGCTAGTGGCAGAGATCAATGGGAAATTTTC	366
QY	301	CTAGAACATGTGAAACAACTTATAAATCAACAAATAACAGAAATGTAGGAATACGGCA	360
DB	367	CTGAAACATGTAGAACAACTTATAAGACAAACAGTAACAGAAATCTAGGAATACGGCT	426
QY	361	CTTGCTCGATTAACAGTTTATAGGATATCTTTAGAGCCTATCAACAGTCACTTGAAGAT	420
DB	427	ATTGCTCGATTAGAAAGTCTAGGAAGAGGCTATAGATCTTACCAGCAGGCTCTTGAAC	486
QY	421	TGCGTAAACCGTGTATGATGAAGAACGAGAGTGTCTTTATATACCCCAATATAGCC	480
DB	487	TGGTATAGTAACCGAAATGATGAAGATCAAGAGATTAATCTTGAGCGCTATGTTGCT	546
QY	481	TTAGAACTTCGATTTCTTAATGCGATGCCGCTTTTCGCAATTAGAAACCAAGAGTTCCA	540
DB	547	TTAGAACTTCGATTTCTTAATGCGATGCCGCTTTTCGCAATTAGAAACCAAGAGTTCCA	606
QY	541	TTATTAATGTTATGCTCAAGCTGCAAAATTTACACCTATTATTATTTGAGAGATGCCCT	600
DB	607	TTATTAATGTTATGCTCAAGCTGCAAAATTTACACCTATTATTATTTGAGAGATGCCCT	666
QY	601	CTTTTGTGTTAGTGAATTTGGGCTTACATCGCAGGAATTTCAAGTTATTATGAGCGCAA	660
DB	667	CTTTTGTGTTAGTGAATTTGGGCTTACATCGCAGGAATTTCAAGTTATTATGAGCGCAA	726
QY	661	GTGGAACAAACGAGAGATTTTCGACTATTTCGTAATGGTATATACAGGCTTAAAT	720
DB	727	ATCAGATATACAGAGATTTTCGACTATTTCGTAATGGTATATACAGGCTTAAAT	786
QY	721	AGCTTGAGAGGGAACAAATGCGCAAGTTGGTGGCTTATTAATCAATCCGTTAGAGATCTA	780
DB	787	AACTTAAGAGGGAACAAATGCTGAAATTTGGTGGCTTATTAATCAATCCGTTAGAGACCTA	846
QY	781	ACGTTAGGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGACACTCGCACTTATCCA	840
DB	847	ACGTTAGGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGACACTCGCACTTATCCA	906

QY	841	ATAAATACGAGTGTCTAGTTTAAACAGGAAGTTTATACAGCGCAATTTGAGACCAACAGG	900
DB	907	ATCAATACGAGTGTCTAGTTTAAACAGGAAGTTTATACAGATCAATTTGGGAGAACAAAT	966
QY	901	GTAAT-----ATGGCAAGTATCAATTTGCTATTAATTAATTAATGACCTTCGTTTCCGCT	954
DB	967	GCACCTTCAGGATTTGCAAGTACGAATTTGGTTTAATTAATTAATGACCATCGTTTTCTGCC	1026
QY	955	ATAGAGACTCGGTTATCCGAAGCCGCATCTACTTGAATTTCTTAGAACCACTTACAAAT	1014
DB	1027	ATAGAGCTGCCATTTTCAGGCTCCGCATCTACTTGAATTTTCAGAACCACTTACAAAT	1086
QY	1015	TTTAGCACTTTCATCAAGATGAGTGTCTATAGGATATAGACTTACTCGCGGGGGGACACACA	1074
DB	1087	TACAGTGCATCAAGCCGTTGGAGTAGCACTCAACATATGAATTTATTTGGTGGGACATAGG	1146
QY	1075	ATTCAATCTCGGCAATAGGAGGGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1131
DB	1147	CTTAATCTCGGCAATAGGAGGGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1206
QY	1132	ACTTCTATTAATCTCTGAAGATTTATCAATCTCTCTCGAGACGTATATTTGGACTGAATCA	1191
DB	1207	ACTTCTATTAATCTCTGAAGATTTATCAATCTCTCTCGAGACGTATATTTGGACTGAATCA	1266
QY	1192	TATGCGAGGAGTGTCTTATGCGGGAATTTACCTTGAACCTTATCATGTTGCTTCTTCTGTT	1251
DB	1267	AATGCGAGGACAAATAT-----ACTATTTACTACTCTCTGTAATGAGTACCTTGGGCT	1320
QY	1252	AGATTTAATTTTAGGAACTCTCAGAACTATTTTGAAGAGGTACTGCTACTATATAGTCAA	1311
DB	1321	AGATTTAATTTTATAAACCTCTCAGAACTATTTTGAAGAGGTACTGCTACTATATAGTCAA	1380
QY	1312	CCCTATGAGTCACCTGGGCTTCAATTTAAAGATTTAGAAATCTGAAATTTACCAACAGAAACA	1371
DB	1381	CCGTTATCAGGAGTGGGATTTCAATTTATTTGATTCAGAACTGAAATTTACCAACAGAAACA	1440
QY	1372	ACAGAACGCAAAATTTATGAATCATATAGTATAGTATTTCTCACATAGGCTCATTTTCA	1431
DB	1441	ACAGAACGCAAAATTTATGAATCATATAGTATAGTATTTCTCATATAGGACTAATCATATA	1500
QY	1432	CAATCTAGGTTGATGATCCAGTATATTTCTTGACGCGCAGTGTAGTCAGATCGTACAAT	1491
DB	1501	GGAACACATTTTGAGAGCACCACTTATTTTGACGCGCAGTGTAGTCAGATCGTACAAT	1560
QY	1492	ACCATTTAGTTTACATAGCATAACAAATACCAATTTGTTAAATCATTTCAACCTTTAAATTTCA	1551
DB	1561	ACGATTTGACCAATTAAGATTTACAAATTTGTTAAATCATTTGTTAAATCATTTCAATTTCA	1620
QY	1552	GGTACCTCTGTAGTGGCCAGGATTTTACAGGAGGGGATATTAATTCGAACTTAACGTT	1611
DB	1621	GGTGTACTTGTGTGGAGGCGCAGGATTTTACAGGTTGGGATATCTTCTCGTAGAACAAAT	1680
QY	1612	AATGTTAGTGTACTAGTATGGTCTTAATTTTAAATTAATCATCATTTACAGCGGTATCGC	1671
DB	1681	ACGGTACATTTTGAGATATACGATTTAAATTAATTTATTTGTCCTATTTCCCAAGATATCGC	1740
QY	1672	GTGAGAGTGTGTTATGCTCTCTCAAACTGCTGAGGGTAACTGTCGAGGGGAGT	1731
DB	1741	GTAAGGATTCGTTATGCTTCTACTACAGATTTTCAATTTTTCAGGAAATTAATTTGAAC	1800
QY	1732	ACTTACTTTTGTATCAAGGATTTCCCTAGTACTATGAGTGCATAATGAGTCTTTGACATCTCAA	1791
DB	1801	ACTGTTAAATTTGGTAAATTTCTCAAGAACCTATGAATAGGGGGGATTAATTTAGATATAGA	1860
QY	1792	TCATTTAGATTTGCAAGATTTCTCTGTAGTATTTAGTGCATCTGGCAGTCAAACTGCTGGA	1851
DB	1861	AGTTTTAGAACTCAGGATTTTAGTACTCTCTTTTAAATTTTAAATGCCCCAAGCACATTC	1920
QY	1852	ATAAGTATAGTAAATTAATGAGGTAGAACAACTTTTCACTTTTCAATTAATTAATTTGATTTCA	1911
DB	1921	ACATGGGTGCTCAGAGTTTTCAAATCAGGAAGTTTATATAGATAGTTCGAAATTTGTT	1980

QY 1912 CCAATTACTGCAACTTTCGAAGCAGAAATACGATTTAGAAAGGGCGCAAGAGGGCGTGAAT 1971
 DB 1981 CCAGCAGAGTAAACATTTGAGGCGAATATGATTTAGAAAGAGCACAACAAAGGGCGTGAAT 2040
 QY 1972 GCTCTGTTTACTAATAGGAATCAAGAGATTGAAACAGATGTGACAGATTATCATATT 2031
 DB 2041 GCTCTGTTTACTTCTACAAATCCAAAGAGATTGAAACAGATGTGACAGATTATCATATT 2100
 QY 2032 GATCAAGTATCCAAATTTAGTGGCGTGTATCGGATGAAATTTCTGCTAGATGAAAGAGA 2091
 DB 2101 GACCAAGTCTCAATATGTTGGCGCATGTTTATCAGATGAATTTTCTGATGAGAAGCGA 2160
 QY 2092 GAATTAATTTGAGAAAGTGAATATGCGAAACGATCTCAGTGAATGAAAGAACTTTACTCCAA 2151
 DB 2161 GAATTAATTTGAGAAAGTGAATATGCGAAACGATCTCAGTGAATGAAAGAACTTTACTCCAA 2220
 QY 2152 GATCCAAACTTTCATCCATCAATTAAGCAACGAGACTTCAATCTCTAATAGAGCAATCG 2211
 DB 2221 GATCCAAACTTTCATCCATCAATTAAGCAACGAGACTTCAATCTCTAATAGAGCAATCA 2280
 QY 2212 AATTTCCATCTTATCCATGAACAATCTGAACATGGAATGTTGGGGAAGTGAGACATTTACA 2271
 DB 2281 AACTTTCCCTCTTATTAATGAGCTATCTGAACATGGAATGTTGGGGAAGTGAGACATTTACC 2340
 QY 2272 ATCCAGGAAGGAATGACGTATTTAAAGAGAAATAGCTCACACTACCGGGGACTTTTAAT 2331
 DB 2341 ATTCCAGGAAGGAATGACGTATTTAAAGAGAAATAGCTCACACTACCGGGGACTTTTAAT 2400
 QY 2332 GAGTGTATTCGACGTATTTATCAAAAATAGGAGAGTCCGAAATTAAGCTTATATCT 2391
 DB 2401 GAGTGTATTCGAAATTTATTAATCAAAAATAGGAGAGTCCGAAATTAAGCTTATATCT 2460
 QY 2392 CGCTACCAATTAGAGGGTATTTGAAGATAGTCAAGATTTAGAGATATTTGATTCGT 2451
 DB 2461 CGCTATCAATTAAGAGGGTATTTGAAGATAGTCAAGATTTAGAGATTTATTTAAATCGT 2520
 QY 2452 TATAATGCGAAAATCAAAACATTTGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTTCA 2511
 DB 2521 TACAATGCAAGCATGAACATTTGATGTTCCAGGTACCGATTCCTATGCGCGCTTTTCA 2580
 QY 2512 GTTGAAAGCCCAATCGGAAGTGGCGAGAACCGAATCGATCGCACCAATTTGGAATGG 2571
 DB 2581 GTTGAAAGCCCAATCGGAAGTGGCGAGAACCGAATCGATCGCACCAATTTGGAATGG 2640
 QY 2572 AATCTGATCTAGATTTCTTCGACAGATGGAGAAAATGTCGCGATCATTTCCCATCAT 2631
 DB 2641 AATCTGATCTAGATTTCTTCGACAGATGGAGAAAATGTCGCGATCATTTCCCATCAT 2700
 QY 2632 TTCTCTTTGGATATTGATATTGGATGCACAGACTTGCATGAGAAATCTAGGCGTGTGGGTG 2691
 DB 2701 TTCATTTGGATATTGATTTGGGTGCACAGACTTGCATGAGAACTTAGGCGTGTGGGTG 2760
 QY 2692 GTATTCAGATTAAGACGCAAGAGTCTATGCAAGACTAGGGAATCTGGAAATTTATGAA 2751
 DB 2761 GTATTCAGATTAAGACGCAAGAGTCTATGCAAGACTAGGGAATCTGGAAATTTATGCA 2820
 QY 2752 GAGAAACCATTTATAGGAGACACTGCTCGTGTGCAAGAGCAGAGAAATGAGAAATCGAG 2811
 DB 2821 GAGAAACCATTTATAGGAGACACTGCTCGTGTGCAAGAGCAGGAAAATGAGAAATGAG 2880
 QY 2812 GACAAACGTTGAAAACCTCAATTTGAAAACAAAACGAGTATATACAGAGGCAAAAGAGCT 2871
 DB 2881 GACAAACGTTGAAAACCTCAATTTGAAAACAAAACGAGTATATACAGAGGCAAAAGAGCT 2940
 QY 2872 GTGGATGCTTTATTTGTAGATTTCTCAATATAATATAGATTAACAGCGGATCAAAACATGGC 2931
 DB 2941 GTGGATGCTTTATTTGTAGATTTCTCAATATAATATAGATTAACAGCGGATCAAAACATGGC 3000
 QY 2932 ATGATTCATCGCGCAGATTAACCTTTGTCGAAATTCGAGAGCTTATCTGTGAGAAATTA 2991
 DB 3001 ATGATTCATCGCGCAGATTAACCTTTGTCGAAATTCGAGAGCTTATCTTTGAGAAATTA 3060
 QY 2992 TCTGTTATCCCGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAGGTGCGATTTATCACT 3051

DB 3061 CCTGTTATCCAGGTGTAAATCGGAAATTTTGAAGAAATAGAAAGTCACTATCACT 3120
 QY 3052 GCAATCTCCCTATACGATCGGAGAAATGTCGTTAAAAATGGTCAATTTAATATGATTA 3111
 DB 3121 GCAATGCTCTTATACGATCGGAGAAATGTCGTTAAAAATGGTCAATTTAATATGATTA 3180
 QY 3112 GCATGCTCGAAATGTTAAAAAGGCATGTAGATGTACAACAGAGCCATCACCGCTTCTGTCCCT 3171
 DB 3181 ACATGTTGAAATGTTAAAAAGGCATGTAGATGTACAACAGAGCCATCATCTTCTGACCTT 3240
 QY 3172 GTTATCCCAAGATGGGAAGCAAGATGTCAAGCAAGTTTCCGCTCTGTCCGGGCGCTGGC 3231
 DB 3241 GTTATCCCAAGATGGGAAGCAAGATGTCAAGCAAGTTTCCGCTCTGTCCGGGCGCTGGC 3300
 QY 3232 TATATCTCTCCGCTCACAGGTACAAAGAGGATATGGAGAGGGTGTGTACGATCCAT 3291
 DB 3301 TATATCTCTCCGCTCACAGGTACAAAGAGGATATGGAGAGGGTGTGTACGATCCAT 3360
 QY 3292 GAAATCGAAACAAATACAGACGAACTTAAAAATTTAAAAAATCTGTGAAGAGAGGATGTAT 3351
 DB 3361 GAAATCGAAACAAATACAGACGAACTTAAAAATTTAAAAAATCTGTGAAGAGAGGATGTAT 3420
 QY 3352 CCAACGATACAGGAACTGTATTAATGATTAATCTGCAACCAAGGTACAGAGTATGTAAT 3411
 DB 3421 CCAACGATACAGGAACTGTATTAATGATTAATCTGCAACCAAGGTACAGAGTATGTAAT 3480
 QY 3412 TCCCGTAAATGCTGGATATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 3471
 DB 3481 TCCCGTAAATGCTGGATATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 3540
 QY 3472 AAACCGACTTATGAAGAAACGCTATACAGATGTACGAAGAGATTAATCATTTGTAATAT 3531
 DB 3541 AAACCGACTTATGAAGAAACGCTATACAGATGTACGAAGAGATTAATCATTTGTAATAT 3600
 QY 3532 GACAGAGGTATGTGAATTTATCCACATCAACAGCTGTTATATGACAAAAGAAATTAGAA 3591
 DB 3601 GACAGAGGTATGTGAATTTATCCACAGTACCCAGTCCAGCTGTTATGTGACAAAAGAAATTAGAA 3660
 QY 3592 TACTTCCCAAGAACCGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 3651
 DB 3661 TACTTCCCAAGAACCGATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 3720
 QY 3652 GTAGACAGCGTGGAAATTTACTCTCTTATGGAGGAATAG 3687
 DB 3721 GTAGATAGCGTGGAACTACTCTCTCATGGAGAAATAG 3756

RESULT 11
 AAV83927
 ID AAV83927 standard; DNA; 3934 BP.
 XX
 AC AAV83927;
 XX
 DT 03-MAR-1999 (first entry)
 XX
 DE DNA encoding a CryET5 protein.
 XX
 KW CryET4; Bacillus thuringiensis strain EG5847; crystal toxin; CryET5;
 Lepidoptera; Helicoverpa zea; resistant strain; Plutella xylostella;
 Spodoptera exigua; S. frugiperda; S. frugiperda; Trichoplusia ni; ds.
 XX
 OS Bacillus thuringiensis.
 XX
 Key Location/Qualifiers
 CDS 67..3756
 FT /*tag= a
 FT /product= "CrtET5"
 XX
 PN US5854053-A.
 XX 29-DEC-1998.
 PD
 XX

XX insecticidal protein; plant; pesticide; gene therapy;
KW lepidopteran insect pest; transgenic plant;
KW insect infestation resistance; monocot; dicot; cry1Bb; gene; ds.
XX

OS Bacillus thuringiensis.

XX Key Location/Qualifiers
FH 1.3687
FT /tag= a
FT /product= "B thuringiensis cry1Bb"
FT /partial
FT /note= "No start or stop codon"

XX WO2004020636-A1.

XX 11-MAR-2004.

XX 26-AUG-2003; 2003WO-US026510.

XX 29-AUG-2002; 2002US-0407428P.

XX (MONS) MONSANTO TECHNOLOGY LLC.
PA (BOGD/) BOGDANOVA N N.
PA (ROMA/) ROMANO C P.

XX Bogdanova NN, Romano CP;

XX WPI: 2004-269221/25.

XX P-PSDB; ADR98479.

XX New polynucleotide sequence optimized for expression of an insecticidal protein in a plant, useful in the control of Lepidoptera insect pests, and for producing transgenic plants with the ability to resist insect infestations.

XX Disclosure; SEQ ID NO 1; 138pp; English.

XX This invention relates to a novel polynucleotide sequence optimised for expression of an insecticidal protein in a plant. The invention may be useful for the production of pesticides while the disclosed sequences may be used for gene therapy. The polynucleotide sequence and methods are useful in the control of lepidopteran insect pests, and for producing transgenic plants with the ability to resist infestations. The invention provides polynucleotide sequences with enhanced, improved and optimised expression in monocot and dicot plant species. The present sequence is that of the (partial) B thuringiensis cry1Bb gene which is related to the invention.

XX Sequence 3687 BP; 1248 A; 621 C; 796 G; 1022 T; 0 U; 0 Other;

Query Match 72.8%; Score 2684.2; DB 12; Length 3687;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 3085; Conservative 0; Mismatches 593; Indels 15; Gaps 3;

QY 1 TTGACTTCAAAATAGGAAAATGAGAAATGAAATATATAATGCTTTATCGAATCCAGCTGTA 60

DB 1 TTGACTTCAAAATAGGAAAATGAGAAATGAAATATATAATGCTTTATCGAATCCAGCTGTA 60

QY 61 TCGAATCATTCACAAATGGAATCTATACACAGATGCTGTAATGAGAAATCTTTGTGT 120

DB 61 TCGAATCTTCCACGCAAAATGAATCTATACACAGATGCTGTAATGGAATAGCTTTGTGT 120

QY 121 ATAGCCGAGGGGAATATCAATCCACTGTTAGCGCATCAACAGTCCAAACGGGTATT 180

DB 121 GTAGCCGAGGTGAACAAATATGTATTCATTTGTTAGCGCATCAACAGTCCAAACGGGTATA 180

QY 181 AACATAGCTGCTAGAACTACTAGTGTATTAAGCGGTACCGTTCTGACAAAATAGCTAGT 240

DB 181 AACATAGCTGCTAGAAATATTTGGCGGTATTAAGTGTGCGGTTCTGCTGGACAACTAGCTAGT 240

QY 241 TTTTATAGTTTCTTGTGTGTAATTATGCGCCCGCGGACAGATCAGTGGGAAATTTTC 300

DB TTTTATAGTTTCTTGTGTGTAATTATGCGCCCGCGGACAGATCAGTGGGAAATTTTC 300

Db	241	TTTTATAGTTTCTTGTGTGGAATTTATGCGCTAGTGGCAGAGATCCATGGGAAATTTTC	300
QY	301	CTAGAACATGTGCAACAACTTATAAATCAAAATAACAGAAAAATGCTTAGGAATACGGCA	360
Db	301	CTGGAACATGTAGAACAACTTATAAGACAAACAAAGTAACAGAAAAATAGTAGGAATACGGCT	360
QY	361	CTTGCTCGAATACAAAGGTTTAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTGAAGAT	420
Db	361	ATTGCTCGAATTAGAAGGCTCTAGGAAGAGGCTATAGATCTTACCAGCAGGCTCTTGAAC	420
QY	421	TGGCTAGAAAAACCGTCATGATGCAAGAACGAGAAGTGTCTTTTATACCCAAATATATAGCC	480
Db	421	TGGTTAGATAACCGAATGATGCAAGATCAAGAGCAATTTCTTGGAGCGCTATGTGCT	480
QY	481	TTAGAACTTTGATTTTCTTAATGCGATGCGCTTTTCGCAATTAGAGAAACCAAGATTCCA	540
Db	481	TTAGAACTTTGATTTTCTTAATGCGATGCGCTTTTCGCAATTAGAGAAACCAAGATTCCA	540
QY	541	TTAATTAATGTTATGCTCAAGCTGCAAAATTTTACACCTATTTATTTAGAGACGCTCT	600
Db	541	TTAATTAATGTTATGCTCAAGCTGCAAAATTTTACACCTATTTATTTAGAGACGCTCT	600
QY	601	CTTTTGTGTTAGTGAATTTTGGGCTTACATCGCAGGAAATTTCAACGTTTATTTATGAGCCCAA	660
Db	601	CTTTTGTGTTAGTGAATTTTGGGCTTACATCGCAGGAAATTTTCCGATGTTAACTAATTTACCAAGAACAA	660
QY	661	GTGGAACAAACGAGAGATTTATTCGCACTATTCGCTAGAAATGTTATTAATCAAGTCTAAAT	720
Db	661	ATCAGATATACAGAGAGATTTATTCGCACTATTCGCTAGAAATGTTATTAATCAAGTCTAAAT	720
QY	721	ACGTTGAGAGGACAAATGCGGCAAGTTGGGTCGTTATAATCAANTTCGTAGAGATCTA	780
Db	721	AACTTTAAGAGGACAAATGCTGAAAGTTGCGGTATTAATCAANTTCGTAGAGACCTTA	780
QY	781	ACGTTAGGCGTTATGATCTAGTGGCACTATTTCCCAAGCTATGACACTCGCACTTATCCA	840
Db	781	ACGTTAGGCGTTATGATCTAGTGGCGCTATTTAGGCCCTATTTCCCAAGCTATGATCTGCACTTATCCA	840
QY	841	ATAAATACAGAGTCTCAGTTAAACAGGGAAGTTTATACAGACGCAATTCGAGCAACAGGG	900
Db	841	ATCAATACAGAGTCTCAGTTAAACAGGGAAGTTTATACAGATCCCAATTCGAGCAACAAAT	900
QY	901	GTAAT-----ATGCAAGTATGAATGCTGTAATAATAAATGACACTTCGTTTCGGCT	954
Db	901	GCACCTTCAGAGATTTGCAAGTACGAATGCGTTTAAATAAATGACACTTCGTTTCGGCT	960
QY	955	ATAGAGACTCGGTTATCGAAGCCGCACTACTTGTATTTTCTAGAACAACTTACAAAT	1014
Db	961	ATAGAGGCTGCCATTTTCAGGCGCTTCGCACTACTTGTATTTTCCAGAACAACTTACAAAT	1020
QY	1015	TTTAGCATTCTACGATGAGTGTCTACTAGGCATATGACTTACTGCGGGGGGCACACA	1074
Db	1021	TACAGTGCATCAAGCGTTGGAGTAGCACTCAACATATGAATTTTGGTGGGACATAGG	1080
QY	1075	ATTCAATCTCGGCCAATAGGAGCGGATTAATACCTCAACGCAATGG---GTCTACCAAT	1131
Db	1081	CTTAACCTTCGCCCAATAGGAGGCAATTAATAATACCTCAACCAAGGACTTACTATAAT	1140
QY	1132	ACTTCTATTAACTCTGATGATTAATCATTTCTCTCGAGAGCTATATTGGACTGAATCA	1191
Db	1141	ACTTCAATTAATCTGATGATTAATCATTTCTCTCGAGAGCTTTTATAGACAGATCA	1200
QY	1192	TATGAGAGGAGTCTTCTATGGGGAATTTTACCTTGAACCTTATCATGGTGTCCCTACTGTT	1251
Db	1201	AATGAGGAGCAAAATAT-----ACTATTACTCTCTGTAATGAGTACCTTGGGCT	1254
QY	1252	AGATTTAATTTTAGGAACCCCTCAGAAATATCTTTTGAAGAGGTACTGCTTAATATAGTCAA	1311
Db	1255	AGATTTAATTTTATAAACCCCTCAGAAATTTTATGAAGAGGCGGCCACTACTACAGTCAA	1314
QY	1312	CCCTATGAGTCACTGGGCTTCAATTTAAAGATTGAGAACTGAATTTACCCAGAAACA	1371
Db	1315	CCGTATCAGGAGTGTGGGATTTCAATTTATTTGATTTTCAGAACTGAATTTACCCAGAAACA	1374

1372 ACAGAACGACCAAAATATAGTAATCATATAGTATCATAGGTTATCTCACATAGGGCTCAITTTCA 1431
1375 ACAGAACGACCAAAATATAGTAATCATATAGTATCATATAGTATCTATATAGACATTAATCATTA 1434
1432 CAATCTAGGTGATGATGACAGTATATCTTGTGACGACCGTATAGTACAGATCGTACAAAT 1491
1435 GGAACACACTTTGAGAGCACCAGTCTATCTTGTGACGCACTCGTAGTCAGATCGTACGAAT 1494
1492 ACCATTAGTTTCAGATAGCATAACACAAATACCAATCGGTAAATCATTTCAACTTTAATTTCA 1551
1495 ACGATTGGAACCAATAGAAATTAACAAATACCAATCGGTAAAGCACTGAAATCTTCAATTTCA 1554
1552 GGTACCTCTCTAGTCTAGTGGCCAGGATTTACAGGAGGGATATAATCCGAACTAACGTTT 1611
1555 GGTGTTACTGTTGTGGAGGGCCAGGATTTACAGGTGGGATATCTTGTGAGAACAAAT 1614
1612 AATGGTAGTGTACTAAGTATGGTCTTAATTTTAAATAATACATATACAGCGGTATCGC 1671
1615 ACGGTACATTTGGAGATATACGATTAAATATTAATGTGCCATTATCCCAAGATATCGC 1674
1672 GTGAGAGTTGTTATGCTGCTTCTCAACAAATGGTCTGAGGTAATCTGTCGAGGGAGT 1731
1675 GTAAGGATTCGTTATGCTTCTACTACAGATTTACAATTTTTTACAGAAATTAATGGAACC 1734
1732 ACTACTTTTGTATCAAGGATTTCCCTAGTACTATGAGTGCAAAATGAGTCTTTGACATCTCAA 1791
1735 ACTGTTAAATTTGGTAATTTCTCAAGAACTAATGAATAGGGGGATTAATTTAGNAATATAGA 1794
1792 TCATTTAGATTTCCAGAAATTTCTGTAGGTATTTAGTGTGCACTGGGAGTCAAACTGCTGA 1851
1795 AGTTTTAGAACTTCAGGATTTAGTACTCTCTTTTAAATTTTAAATGCCCAAGACATTC 1854
1852 ATAACTATAAGTAATTAATGCAGGTAGACAAACGTTTTCACTTTGATATAAATTTGAAATTT 1911
1855 ACATGGGTGCTCAGAGTTTTTCAATCAGGAAGTTTATATAGATAGTTCGAATTTGTT 1914
1912 CCAATTACTGCAACCTTCGAAGCAGAAATACGATTTAGAAAGGGCGCAGAGCGGTGAAT 1971
1915 CCAGCAGAGTAACATTTGAGGCAGATATGATTTAGAAAGAGCACAAGAGCGGTGAAT 1974
1972 GCTCTGTTTACTAATACGAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAAT 2031
1975 GCTCTGTTTACTTCAAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAAT 2034
2032 GATCAAGTATCCAAATTTAGTGGGCTGTTTATCGATGGAATTTCTGCTAGATGAAAGAGA 2091
2035 GACCAAGTGTCCAAATTTAGTGGGCTGTTTATCGATGGAATTTTCTGCTAGATGAAAGAGA 2094
2092 GAATTAATTTGAGAAATGAAATATGCAAAACGACTCAGTGATGAAAGAACTTTACTCCAA 2151
2095 GAATTAATTTGAGAAATGAAATATGCAAAACGACTCAGTGATGAAAGAACTTTACTCCAA 2154
2152 GATCCTAACTTACATCTCAATTAAGAAACGAGCTTCAATCTACTAATGAGCAATCG 2211
2155 GATCCTAACTTCAATCTCAATCTCAATTAAGAAATTTAGGTTTCCGATCCATCGACAAATCA 2214
2212 AATTTTCAATCTATCAATGAAATCTGAAATCTGAAATCTGAAATCTGAAATCTGAAATCTGAA 2271
2215 AACTTCCCTCTATTAATGAGCTATCTGAAATCTGAAATCTGAAATCTGAAATCTGAAATCTGAA 2274
2272 ATCCAGAAAGGAATGACGTATTTAAAGAAATTTAGTCCATCTACACTACCGGGGCTTTTAAAT 2331
2275 ATTCAGAAAGGAATGACGTATTTAAAGAAATTTAGTCCATCTACACTACCGGGGCTTTTAAAT 2334
2332 GAGTGTATCCGAGCTATTTATCAAAATTAAGAGAGTCCGAATTTAAAGCTTATACT 2391
2335 GAGTGTATCCAAATTTATTAATCAAAATTAAGAGAGTCCGAATTTAAAGCTTATACT 2394
2392 CGCTACCAATTAAGAGGTATATTAAGATAGTCAAGATTTAGAGATATATTTGATTCGT 2451
2395 CGCTATCAATTAAGAGGTATATTAAGATAGTCAAGATTTAGAGATATATTTTAAATTCGT 2454

2452 TATAATGCGAAACATGAACATTTGGATGTTTCCAGGTACCGAGTCCGATGCGCGCTTTCA 2511
2455 TACATGCAAGCAGTGAACATTTGGATGTTTCCAGGTACCGATTCCTATGCGCGCTTTCA 2514
2512 GTTGAAGCCCAATCGGAAGGTGCGGAGAACCGAATCGATCGCACCAATTTTGAATCG 2571
2515 GTTGAAGCCCAATCGGAAGGTGCGGAGAACCGAATCGATCGCACCAATTTTGAATCG 2574
2572 AATCCTGATCTAGATGTTTCTGCGAGATGAGAAAGATGTCGCGATCATTTCCCATCAT 2631
2575 AATCCTGATCTAGATGTTTCTGCGAGATGAGAAAGATGTCGCGATCATTTCCCATCAT 2634
2632 TTTCTTTTGGATTAATTTGGATGTCACAGACTTTCGATGAGAACTTAGCGCTGTGGGTG 2691
2635 TTTCACTTTTGGATTAATTTGGGTGCGAGACTTTCGATGAGAACTTAGCGCTGTGGGTG 2694
2692 GTATTCAAGATTAAGACGCGAGAAAGGTCTCATGCAAGACTAGGAATCTTGAATTTTAA 2751
2695 GTATTCAAGATTAAGACGCGAGAAAGGTCTCATGCAAGACTAGGAATCTTGAATTTTAA 2754
2752 GAGAAACCAATTAATAGGAGAAAGCTGCTGCTGTTGAGAGAGCAGAGAAATTTGAGAA 2811
2755 GAGAAACCAATTAATAGGAGAAAGCTGCTGCTGTTGAGAGAGCAGAGAAATTTGAGAA 2814
2812 GACAAACGTTGAAACCTACAAATTTGAAACCAAAACGAGTATATACAGAGCAGAAAGAGCT 2871
2815 GACAAACGTTGAAACCTACAAATTTGAAACCAAAACGAGTATATACAGAGCAGAAAGAGCT 2874
2872 GTGGATGCTTTAATTTGTAGATTTCTCAATATATAGATTAACAGCGGATACAAACATTTGGC 2931
2875 GTGGATGCTTTAATTTGTAGATTTCTCAATATATGATCAATTTCAAGCGGATACAAACATTTGGC 2934
2932 ATGATTCATGCGGAGATAACTTCTGCTGATCGAAATTCGAGAGCTTATCTGTCAGAAATTA 2991
2935 ATGATTCATGCGGAGATAACTTCTGCTGATCGAAATTCGAGAGCTTATCTTTCAGAAATTA 2994
2992 TCTGTTATCCCGGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAGGTGCGCATTTACT 3051
2995 CCGTTATCCCGGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAGGTGCGCATTTACT 3054
3052 GCATCTCTCTATACGATCGGAGAAATGTCGTTTAAATTTGATGATTTTAAATGATTA 3111
3055 GCATCTCTCTATACGATCGGAGAAATGTCGTTTAAATTTGATGATTTTAAATGATTA 3114
3112 GCATCTCTCTATACGATCGGAGAAATGTCGTTTAAATTTGATGATTTTAAATGATTA 3171
3115 ACATGTTGGAATGTTAAAGGCAATGATGATTAACAGAGCCATCATGTTCTGACCTT 3174
3172 GTTATCCAGAAATGGAAGCAGAAAGTGTCAACAGCAGTTTCGCTGTCGCGGGCGTGGC 3231
3175 GTTATCCAGAAATGGAAGCAGAAAGTGTCAACAGCAGTTTCGCTGTCGCGGGCGTGGC 3234
3232 TATATCTCTCTGTCACAGCGTACAAAGGGATGAGAGGGTGTGTAACGATTCAT 3291
3235 TATATCTCTCTGTCACAGCGTACAAAGGGATGAGAGGGTGTGTAACGATTCAT 3294
3292 GAAATCGAGAAACAAATACAGAGAACTTAAATTTTAAATTTGTAAGAGAGAGGATGAT 3351
3295 GAAATCGAGAAACAAATACAGAGAACTTAAATTTTAAATTTGTAAGAGAGAGGATGAT 3354
3352 CCAACGGATACAGGAACTGTAATGATTTACTGCAACCAAGGTTACAGAGCATGTAAT 3411
3355 CCAACGGATACAGGAACTGTAATGATTTACTGCAACCAAGGTTACAGAGCATGTAAT 3414
3412 TCCCGTAAATGCTGATGAGATGATGAGTGTGATGATGATGATGATGATGATGATGATGAT 3471
3415 TCCCGTAAATGCTGATGAGATGATGAGTGTGATGATGATGATGATGATGATGATGATGAT 3474
3472 AAACCGACTTATGAAGAGAAACGATATACAGATGTAACGAGATTAATCTGTTGAAAT 3531
3475 AAACCGACTTATGAAGAGAAACGATATACAGATGTAACGAGATTAATCTGTTGAAAT 3534
3532 GACAGAGGGTATGTAATTTATCCACCACTACCGCTGTTTATGACAAAAAGAAATTAGAA 3591

```

Db 3535 GACAGAGGGTATGTGAATTATCCACAGTACCAGCTGGTTATGTGACAAAAGAAATTAGAA 3594
QY 3592 TACTTCCCGAAGAACCGAATAGGTATGGATTGGAGATTGGAGAAACGGAAGGAAATTTATT 3651
Db 3595 TACTTCCCGAAGAACAGATACAGTATGGATTGGAGATTGGAGAAACGGAAGGAAATTTATT 3654
QY 3652 GTAGACAGCGTGAATTAATCTCTTATGGAGAA 3684
Db 3655 GTAGATAGCGTGAATTAATCTCTTATGGAGAA 3687

RESULT 13
ID AAS04854
AA AAS04854 standard; DNA; 3627 BP.
AC AAS04854;
XX
XX
DT 07-SEP-2001 (first entry)
XX
DE Bacillus thuringiensis DNA encoding partial mutant CryIbA.
KW Crystal protein; CryIIa; CryIbA; moth; butterfly; Colorado potato beetle;
KW db; mutant.
XX
OS Bacillus thuringiensis.
XX
XX Location/Qualifiers
FH Key 1..3627
FT CDS /*tag= a
FT /*product= "CryIbA"
FT /*partial
FT /*note= "No stop codon"
FT sig_peptide 1..57
FT /*tag= b
FT mat_peptide 58..3651
FT /*tag= c
FT mutation /*label= Mature CryIbA
FT /*tag= d
FT mutation /*tag= e
FT /*tag= e

EP1099760-A1.
XX
XX 16-MAY-2001.
XX
XX 09-NOV-1999; 99EP-00203723.
XX
XX 09-NOV-1999; 99EP-00203723.
XX (CPRO-) CPRO-DLO CENT PLANTENVERDELINGS REPROD.
XX
XX De Maagd RA, Bosch HJ;
XX
XX WPI; 2001-337141/36.
XX
XX P-PSDB; AAU02094.
XX
XX New hybrid Bacillus thuringiensis hybrid toxins comprising structural
XX domains derived from at least 2 different crystal proteins, such as
XX CryIIa and CryIbA, and having insecticidal activity, useful for combating
XX insects.
XX
XX Example; Page 22-23; 43pp; English.
XX
XX The sequence encodes B. thuringiensis (Bt) crystal protein CryIbA,
XX mutated to allow cloning of domain III or domains I and II, to make the
XX hybrid protoxins of the invention. The hybrid toxins of the invention,
XX having structural domains I, II and III in this order starting from the N
XX -terminal derived from at least 2 different crystal proteins, are useful
XX for protecting plants against pest insects, e.g. moths, butterflies and
XX Colorado potato beetle or for combating insects

```

```

SQ Sequence 3627 BP; 1184 A; 648 C; 820 G; 975 T; 0 U; 0 Other;
Query Match 71.3%; Score 2630.2; DB 4; Length 3627;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 3081; Conservative 0; Mismatches 533; Indels 81; Gaps 5;
QY 2 TGACTTCAATAGGAAATAGAAATGAAATATAAATGCTTTATCGATTCCAGCTGTAT 61
Db 2 TGACTTCAATAGGAAATAGAAATGAAATATAAATGCTTTATCGATTCCAGCTGTAT 46
QY 62 CGAATCATTTCCACACAAATGGATCTATCACAGATGCTCGTATTGAGGATTTCTTGTA 121
Db 47 CGAATCATTTCCGACAAATGGATCTATTAACAGATGCTCGTATTGAGGATTTCTTGTA 106
QY 122 TAGCCGAGGGGAATATATCAATCCACTTTAGCGCATCAACAGTCCAAAACGGGTATTA 181
Db 107 TAGCCGAGGGGAACAATATCGATCCATTTGTTAGCGCATCAACAGTCCAAAACGGGTATTA 166
QY 182 ACATAGCTGGTAGNATCTAGGTGTTATTAGGCGTACCGTTTGTGACAAATAGCTAGTT 241
Db 167 ACATAGCTGGTAGNATCTAGGTGTTATTAGGCGTACCGTTTGTGACAAATAGCTAGTT 226
QY 242 TTATATAGTTTCTTGTGTTGTAATTAATGCGCCCGCGCAGAGATCAGTGGAAATTTTCC 301
Db 227 TTTATAGTTTCTTGTGTTGTAATTAATGCGCCCGCGCAGAGATCAGTGGAAATTTTCC 286
QY 302 TAGAACATGTCGAACAACTTATAAATCAACAAATAACAGAAATGCTAGGAATACGGCAC 361
Db 287 TAGAACATGTCGAACAACTTATAAATCAACAAATAACAGAAATGCTAGGAATACGGCAC 346
QY 362 TTGCTCGATTACAGGTTTAGGAGATTCCTTTAGACCTTCAACAGTCACTTGAAGATT 421
Db 347 TTGCTCGATTACAGGTTTAGGAGATTCCTTTAGACCTTCAACAGTCACTTGAAGATT 406
QY 422 GCTAGAAAACCGTGATGATGCAAGAACGAGAGTGTCTTTATATACCAATATATAGCCT 481
Db 407 GCTAGAAAACCGTGATGATGCAAGAACGAGAGTGTCTTTATATACCAATATATAGCCT 466
QY 482 TAGAACTTGATTTCCTTAATGCGATGCCGCTTTTCGCAATTAGAAAACCAAGAGTTCCAT 541
Db 467 TAGAACTTGATTTCCTTAATGCGATGCCGCTTTTCGCAATTAGAAAACCAAGAGTTCCAT 526
QY 542 TATTAATGGTATATGCTCAAGCTGCAAAATTAACCTATATATTTAGAGAGATGCCCTC 601
Db 527 TATTAATGGTATATGCTCAAGCTGCAAAATTAACCTATATATTTAGAGAGATGCCCTC 586
QY 602 TTTTTCGTAGTGAATTTGGGCTTTACATCGCAGAAATTCACCGTATATATGAGCGCAAG 661
Db 587 TTTTTCGTAGTGAATTTGGGCTTTACATCGCAGAAATTCACCGTATATATGAGCGCAAG 646
QY 662 TGGACAAAACGAGAGATTTCCGACTATTTGCGTAGAATGGTATAATACAGGCTCTAAATA 721
Db 647 TGGACAAAACGAGAGATTTCCGACTATTTGCGTAGAATGGTATAATACAGGCTCTAAATA 706
QY 722 GCTTAGAGGGAACAATGCCGCAAGTTGGTGCTTATTAATCAATTCGTTAGAGATCTAA 781
Db 707 GCTTAGAGGGAACAATGCCGCAAGTTGGTGCTTATTAATCAATTCGTTAGAGATCTAA 766
QY 782 CGTTAGGCTTATAGTCTAGTGGCACTATTCGAGCTATGACCTCGCACATTATCCAA 841
Db 767 CGTTAGGCTTATAGTCTAGTGGCACTATTCGAGCTATGACCTCGCACATTATCCAA 826
QY 842 TAAATACGAGTGTCTCAGTTTAAACAGGGAAGTTTATACAGACGCAATTTGGAGCAACAGGGG 901
Db 827 TAAATACGAGTGTCTCAGTTTAAACAGGGAAGTTTATACAGACGCAATTTGGAGCAACAGGGG 886
QY 902 TAAATATGCAAGTATGAATTTGGTATAATTAATGCACTTCGTTTTCGCTATATAGAGA 961
Db 887 TAAATATGCAAGTATGAATTTGGTATAATTAATGCACTTCGTTTTCGCTATATAGAGA 946
QY 962 CTGCGGTTATCCGAGCCGCGCATCTACTGATTTTCTAGAACAACTTACAAATTTTAGCA 1021
Db 947 CTGCGGTTATCCGAGCCGCGCATCTACTGATTTTCTAGAACAACTTACAAATTTTAGCG 1006

```

1022 CTTTCATCAGATGGAGTGCTACTAGGCAATATGACTTACTGCGGGGCAACAAATTCAT 1081
1007 CTTTCATCAGATGGAGTAATACTAGGCAATATGACTTACTGCGGGGCAACAAATTCAT 1066
1082 CTCGGCAATAGGAGCGGATTAATACTCAACGCAATGGGTCTACCAATCTCTATTA 1141
1067 CTCGGCAATAGGAGCGGATTAATACTCAACGCAATGGGTCTACCAATCTCTATTA 1126
1142 ATCTGTAAGATATCATCTCTCTCGAGCGTATATGGAATGATCAATATGCGAGG 1201
1127 ATCTGTAAGATATCATCTCTCGAGCGTATATGGAATGATCAATATGCGAGG 1186
1202 TGCTTCTATCGGGAATTTACTCTGAACTATTCATGGTGTCCCTACTGTTAGATTTAAT 1261
1187 TGCTTCTATCGGGAATTTACTCTGAACTATTCATGGTGTCCCTACTGTTAGATTTAAT 1246
1262 TTAGGAACCTCTAGATATCTTTTGAAGAGGTACTGCTAACTATAGTCAACCTATGAGT 1321
1247 TTAGGAACCTCTAGATATTTCTGATAGAGGTACCGCTAACTATAGTCAACCTATGAGT 1306
1322 CACTGGGCTTCAATTAAGATTCAGAACTCAATTTACCACAGAAACAAACAGACGAC 1381
1307 CACTGGGCTTCAATTAAGATTCAGAACTCAATTTACCACAGAAACAAACAGACGAC 1366
1382 CAAATATGATCATATAGTATAGTATCTCACATAGGCTCAATTTCAATCTAGGG 1441
1367 CAAATATGATCATTTACAGTCACAGGTATCTCATATAGTATAATTTTACAATCCAGG 1426
1442 TGCAATGACAGATATTTCTTTGACCGCACGCTAGTGCAGATCTGACAAATCAATAGTT 1501
1427 TGAATGTACCGGTATATTTCTTGGACGCATCTGTAGTGGGACCGTACGAATTCGAC 1486
1502 CAGATAGATAACACAAATACCAATGCTGTAATTCATTTCAACCTTTAATTCAGGTACCTCTG 1561
1487 CAAATAGATCAACCAATCCCAATGGTAAGCATCCGAACTTCTCAAGGTACCACGT 1546
1562 TAGTCAGTCCCGAGATTTACAGGAGGATATAATCCGAACTAACGTTAATGTTAGT 1621
1547 TTGCTAGAGACCAAGATTTACTGTTGGGATATTTCTTGAAGAACGAATCTGTTGGAT 1606
1622 TACTAGTATGGGTCTTAATTTTAATATACATCATACAGCGGTATCGGTGAGAGTTC 1681
1607 TTGCAACCGATAAGAGTAACCTGTTAAGCGACCAATTAACACAAAGATATCGTATAGGATTC 1666
1682 GTTATGCTGCTTCTCAAAACAATGCTCTGAGGGTAACCTGTGCGAGGAGTACTACTTTTG 1741
1667 GCTATGCTTCAACTGATAGATTTGATTTCTTGTATCACTGAGGATCTACTGTHAATA 1726
1742 ATCAAGGATTCCTAGTACTATGAGTGCAATGAGTCTTTGACATCTCAATCAATTTAGAT 1801
1727 ATTTTAGATTCCTACGTACAATGAACAGTGGAGCGAACTAAATACGGAAATTTTGTGA 1786
1802 TTGCAAAATTCCTGTAGGTATAGTATGATCTGGCAGTCAACTGCTGGAATAGATATA 1861
1787 GAGTCTGTTTACTACACCTTTTACTTTTACAAATTCAGATATAATTCGAACGTCTA 1846
1862 GTAATAATGAGGTAGACAAACG- - -TTTCACTTTGATAAAATTTGAATTCATTCCAATTA 1918
1847 TTCAAGGCTTGTAGTGAATGGGAAGTGTATATAGATATAATTTGAATTTTCCGTTA 1906
1919 CTGCAACCTTTGGAAGCAGATACGATTTAGAAAGGCGCAAGGCGGTGAATGCTCTGT 1978
1907 CTGCAACCTTTGGAAGCAGATATGATTTAGAAAGGCGCAAGGCGGTGAATGCTCTGT 1966
1979 TTACTAATACGAATCAAGAGATTCGAAACAGATGTGACATATCATATTTGATCAG 2038
1967 TTACTAATACGAATCAAGAGATTCGAAACAGATGTGACATATCATATTTGATCAG 2026
2039 TATCCAAATTTAGTGGGCTGTTTATCGGATGAATTTCTGCTTAGATGAAAGAGAGATTTAC 2098
2027 TATCCAAATTTAGTGGATTTGTTTATCAGATGAATTTTGTCTGGAATGAAAGAGAGATTTCT 2086

2099 TTTGAGAAATGAAATATGCAAAACGACTCAGTGTGATGAAAGAACTTACTCCAGATCCAA 2158
2087 CCGAGAAATGCAAAATGCAAAACGACTCAGTGTGATGAGCGGAATTTACTTCAAGATCCAA 2146
2159 ACTTCAATCCATCAATTAAGCAACGAGACTTCATATCTACTAATGAGCAATCGAATTTCA 2218
2147 ACTTCAAGAGGATCAATAGCAACAGAC- - - - - 2175
2219 CATCTATCCATGACAACTCTGAACATGATGTTGGGAAGTGAACATTTACATCCAG 2278
2176 - - - - - CTTGGCTGGAGAGGATGACAGATATTTACATCCAG 2212
2279 AAGGAAATCAGCTATTTAAAGAGAAATTTAGCTCACACTACCGGGGACTTTTAATGAGTGT 2338
2213 GAGGAGATCAGCTATTTCAAGAGAAATTTAGCTCACACTACCGGGTACCGTTGATGAGTGT 2272
2339 ATCCGACGTATTTATATCAAAAAATAGATGAGTCAAGACTTTAGAAATCTATTGATCCGTTACAATG 2398
2273 ATCCGACGTATTTATATCAAAAAATAGATGAGTCAAGACTTTAGAAATCTATTGATCCGTTACAATG 2332
2399 AATTAAGAGGATATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTATAATG 2458
2333 AATTAAGAGGATATTTGAAGATAGTCAAGACTTTAGAAATCTATTGATCCGTTACAATG 2392
2459 CGAAACATGAAACATTTGGAATTTCCAGGTACCGAGTCCGTATGCGGCTTTTCAGTTGAAA 2518
2393 CAAACACGAAATAGTAAATGTTGCCAGGCACGGTTCTTTATGCGGCTTTTCAGGCCAAA 2452
2519 GCCCAATCGGAAGGTGCGGAGAACCGAATCGATGCGCACCAATTTGAAATGGAATCTCTG 2578
2453 GTCCAAATCGGAAGGTGCGGAGAACCGAATCGATGCGGCGCACCACTTTGAAATGGAATCTCTG 2512
2579 ATCTAGATTTGTTCTGCGAGATGAGAAAAATGTCGCGCATCATTTCCCATCATTTCTCTT 2638
2513 ATCTAGATTTGTTCTGCGAGACGCGGAAAAATGTCGACATCATTTCCCATCATTTCACT 2572
2639 TGGATATTTGATATTTGATGTCACAGACTTTGATGAGAACTTAGGCGTGTGGGTGATTTCA 2698
2573 TGGATATTTGATTTGATGTCACAGACTTTAAATGAGGACTTAGGTGTATGCGGTGATTTCA 2632
2699 AGATTAAGACCGGAGAGGTCTGCAAGACTAGGAAATCTGGAATTTATTGAAGAGAAC 2758
2633 AGATTAAGACCGGAGAGGTGCGCATGCAAGACTAGGAAATCTGAGTCTTCTGGAAGAGAAC 2692
2759 CATTATTTAGGAGAGCACTGCTCTGCTGTAAGAGACGAGAAAAATGGAAGAGCAAAAC 2818
2693 CATTATTTAGGAGAGCACTGCTCTGCTGTAAGAGACGCGGAGAGAGTGGAGAGACAAAC 2752
2819 GTGAAAACTACAATTTGGAACAAACCGAGTATATACAGAGCAAAAGAGCTGTGGATG 2878
2753 GAGAGAAACGTCAGTTGGAACAAATATTTTATAAAGAGCAAAAGAAATCTGTAGATG 2812
2879 CTTTATTTGTAGATTTCTCAATATATAGTTTACAGCGGATACAAACATTTGCGCATGTT 2938
2813 CTTTATTTGTAACTCTCAATATGATGATTACAAGTGGATACGAACTACCGGATGTT 2872
2939 ATGCGCGAGATAAACTCTGTTCAATCAATTTCCAGAGGCTTATCTGTCAGAAATTTATCTGTTA 2998
2873 ATGCGCGAGATAAAACCGCTTCATAGAAATCCGGGAGCGTATCTGCGAGTGTCTGTGA 2932
2999 TCCCGGTGTAAATCGGAAATTTTGAAGAAATTAAGAGGTGCGCAATTTACTGCAATCT 3058
2933 TTTCCAGGTGTCAATCGGCAATTTTTCGAAGAAATTAGAGGAGCGTATTTTACAGCGTAT 2992
3059 CCTATACGATCGAGAAATGTCGTTTAAATGTTGATTTTAAATGATTTAGATGATGCT 3118
2993 CTTTATATGATGCGGAAATGTCATTTAAATGCGGATTTCAATGATGCTTTATGCT 3052
3119 GGAATGTAAGAGGCGCATGTAGATGTACAA- - - CAGAGCCATCAACCGTTCTGTCTTGTGA 3175
3053 GGAAGTGAAGGTCATGTAGATGTAGAGAGCAAAACCAACCGTTTCGGTTCCTTGTGA 3112
3176 TCCGAAATGGAGACAGAGTGTCAAGACGATTCGCGTCTGTCCGGGCGGTGCTATA 3235

Db 3113 TCCCAAGATGGAGCGCAGAGTGTCAAGAGGTTCTGTCTGTCCAGGTCTGGCTATA 3172
 Qy 3236 TCCTCCGTGTACACGCGTACAAAGAGGATATGGAGGGTGTGTGTACGATCCATGAAA 3295
 Db 3173 TCCCTCGTGTACACGATATTAAGAGGGATATGGAGGGCTGCTAACGATCCATGAGA 3232
 Qy 3296 TCGAGAACATACAGACGAACCTAAATTTAAAACTGTGAAGAGAGGAAGTGTATCCAA 3355
 Db 3233 TCGAAGACATACAGACGAACCTGAAATTCAGCAACTGTGTGAAGAGAGGAAGTGTATCCAA 3292
 Qy 3356 CGGTACAGGACGTGTATGATTTATCTACTGCACACCAAGGTACAGCGATATGT-----A 3409
 Db 3293 ACAACACAGTAACGTGTATTAATTAATCTGGGACTCAAGAGAAATATGAGGTACGTACA 3352
 Qy 3410 ATTCCCGTAATCTCGATATGAGGATGCATATGAAGTTGATACTACAGCATCTGTAAAT 3469
 Db 3353 CTCTCTGTAATCAAGGATATGACGAGCTATGTGTAATACCCCTCCGTACAGCTGAT 3412
 Qy 3470 ACAACCGACTTATGAAGAGAAACGTATACAGATGTACGAGAGATTAATCATTTGTGAAT 3529
 Db 3413 ACGCTTCAGTCTATCAAGAGAAATCGTATACAGATGGACGAGAGAGAAATCCTTTGTGAAT 3472
 Qy 3530 ATGACAGAGGTATCTGAATTTATCCACCACTACCAGCTGTTATATGACAAAGAAATTAG 3589
 Db 3473 CTAACAGAGGCTATGGGATTTACACCACTACCGGCTGGTTATGTATACAAAGAGATTAG 3532
 Qy 3590 AATACTTCCCAAGAACCATTAAGGTATGATGATGAGATGAGAGAAACGGAAGGAAAGTTTA 3649
 Db 3533 AGTACTTCCAGAGACCGATGAGTATGATGAGATGAGATGAGATGAGATGAGATGAGAT 3592
 Qy 3650 TTGTAGACAGCGTGAATTTACTCTCTATGGAGAA 3684
 Db 3593 TCGTGGATAGCGTGAATTTACTCTCTATGGAGAA 3627

RESULT 14
 ID AAS00422 standard; DNA; 3663 BP.
 XX AAS00422;
 AC AAS00422;
 XX 11-MAY-2001 (first entry)
 XX B. thuringiensis Cry1Ba/Cry1Ca (BtC15) hybrid insecticidal toxin DNA.
 XX Insecticidal agent; hybrid toxin; Cry1Ca; Cry1Ba; BtC15; insect control;
 KW pest control; crop; insect resistant plant; transgenic plant;
 KW crystal protein; mutant; ds.
 XX Bacillus thuringiensis.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH 1. 3663
 FT CDS /tag= a
 FT /partial
 FT /product= "BtC15"
 FT /note= "This sequence does not contain a stop codon"
 XX WO200114562-A1.
 XX 01-MAR-2001.
 XX 17-AUG-2000; 2000NO-EP008042.
 XX 19-AUG-1999; 99US-00377511.
 XX (SYNG-) SYNGENTA PARTICIPATIONS AG.
 XX De Maagd RA, Bosch HJ, Carozzi NB, Warren GW;
 XX WPI; 2001-218452/22.

DR P-PSDB; AAU00421.
 XX Novel hybrid insecticidal toxin useful for controlling insects such as
 PT Spodoptera exigua and Plutella xylostella, comprises domains I and II
 PT from Cry1F or Cry1B toxin joined to domain III from Cry1C toxin.
 XX Claim 7; Page 88-89; 99pp; English.

XX The present sequence encodes for 1 of 3 novel hybrid insecticidal toxins
 CC (AAU00419-AAU00422) and is designated BtC15. BtC15 comprises domains I
 CC and II from the Cry1Ba toxin at the N-terminus joined to domain III from
 CC Cry1Ca toxin at the C-terminus. In the 1186 amino acid PFC1 hybrid
 CC sequence, the junction between the Cry1Ba and Cry1Ca toxin domains
 CC corresponds to PFC1 residues 446-454. In the 1221 amino acid BtC13 and
 CC BtC15 hybrid sequences, the junction between the Cry1Ba and Cry1Ca toxin
 CC domains is residues 482-488 in the BtC13 hybrid, and residues 491-494 in
 CC the BtC15 hybrid. The hybrid toxins are constructed by cloning and in
 CC vivo recombination. The hybrid toxins are useful for controlling an
 CC insect such as Spodoptera exigua (beet armyworm), Manduca sexta (tobacco
 CC hornworm), Plutella xylostella (diamondback moth), Ostrinia nubilalis
 CC (European corn borer), Spodoptera frugiperda (fall armyworm), and
 CC Heliothis virescens (tobacco budworm) which are major causes of
 CC vegetable, fruit or ornamental flower crop losses. The nucleotide
 CC sequences encoding the hybrid toxins are useful for producing an insect-
 CC resistant plant by introducing the nucleotide sequence into the plant
 CC preferably a transgenic plant. The nucleotide sequences can also be
 CC expressed in microorganisms such as a virus, bacterium and fungus and the
 CC toxin produced used as an insecticidal agent. Baculoviruses engineered
 CC with the nucleotide sequence encoding a hybrid toxin can be used to
 CC infect insects in vivo resulting in their death. The hybrid toxins have a
 CC greater toxicity level than either of their parent toxins alone

XX Sequence 3663 BP; 1190 A; 650 C; 832 G; 991 T; 0 U; 0 Other;
 SQ Query Match 69.0%; Score 2544.4; DB 4; Length 3663;
 Best Local Similarity 82.1%; Pred. No. 0;
 Matches 3065; Conservative 0; Mismatches 546; Indels 123; Gaps 7;

Qy 2 TGACTTCAAAATAGGAAAAATGAGAAATGAAATATATAATGCTTTATCGATCCAGCTGAT 61
 Db 2 TGGCTTCAATAGGAAAAATGAGAAATGAAATATATAAT-----GCTGAT 46
 Qy 62 CGAATCATTTCCACAAATGATCTATCACAGATGCTCGTATTGAGGATCTTTGTGTA 121
 Db 47 CGAATCATTTCCGCAAAATGATCTATTAACAGATGCTCGTATTGAGGATGATGTA 106
 Qy 122 TAGCCGAGGGGAAATATATCAATCCACTGTGTAGGGATCAACAGTCCAAACGGGTATA 181
 Db 107 TAGCCGAGGGGAAACAAATATCGATCCATTTGTGTAGGCAATCAACAGTCCAAACGGGTATA 166
 Qy 182 ACATAGCTGTGAGAAATACTAGGTGTATTAGGCGTACCGTTTGTGCGACAAATAGCTAGT 241
 Db 167 ACATAGCTGTGAGAAATACTAGGCGTATTGGCGTATGCGTTTGTGCGACAACTAGCTAGT 226
 Qy 242 TTTATAGTTTCTTTGTTGGTGAATTTATGGCCCGCGGAGAGATCAGTGGGAAATTTCC 301
 Db 227 TTTATAGTTTCTTTGTTGGTGAATTTATGGCCCGCGGAGAGATCAGTGGGAAATTTCC 286
 Qy 302 TAGAATCATGTGCAACCACTTATTAATCAACAAATTAACAGAAATGCTAGGAAATAGCGCAC 361
 Db 287 TAGAATCATGTGCAACCACTTATTAATCAACAAATTAACAGAAATGCTAGGAAATAGCGGTC 346
 Qy 362 TTGCTCGAATTACAAGGTTTAGGAGATTCCTTTTAGAGCTCTATCAACAGTCACTTGAAGATT 421
 Db 347 TTGCTCGAATTACAAGGTTTAGGAGATTCCTTCAGAGCTCTATCAACAGTCACTTGAAGATT 406
 Qy 422 GGCTAGAAAACCGTGATGATGCAAGAACGAGAAGTGTCTTTTATACCAATATATAGCCT 481
 Db 407 GGCTAGAAAACCGTGATGATGCAAGAACGAGAAGTGTCTTTTATACCAATATATAGCCT 466
 Qy 482 TAGAATCTGATTTCTTAATCGGATGCGCTTTTCGCAATTTAGAAAACCAAGATTCAT 541
 Db 467 TAGAATCTGATTTCTTAATCGGATGCGCTTTTCGCAATTTAGAAAACCAAGATTCAT 526

Qy 2 TGACTTCAAAATAGGAAATAGGAATGAAATTAATAATGCTTTATCGATTCAGCTGTAT 61
 Db 2 TGGCTTCAAAATAGGAAATAGGAATGAAATTAATAATGCTTTATCGATTCAGCTGTAT 46
 Qy 62 CGAATCAATCCACAAATGATCTATCACAGATGCTCGTATTTGAGGATCTTTTGTGTA 121
 Db 47 CGAATCAATCCGACAAATGATCTATTAACAGATGCTCGTATTTGAGGATGATCTTGTA 106
 Qy 122 TAGCCGAGGGAATAATATCAATCCACTTGTAGCGCATCAACAGTCCAAACGGGTATTA 181
 Db 107 TAGCCGAGGGAACAATATCGATCCATTTGTAGCGCATCAACAGTCCAAACGGGTATTA 166
 Qy 182 ACATAGCTGTAGAAATAGTGTATTAAGCGTACCGTTTGTGGAACAATAGCTAGTT 241
 Db 167 ACATAGCTGTAGAAATAGTGTATTAAGCGTACCGTTTGTGGAACAATAGCTAGTT 226
 Qy 242 TTTATAGTTTCTTTGTTGGTGAATTAATGCGCCCGCGCAGAGATCAGTGGGAATTTTC 301
 Db 227 TTTATAGTTTCTTTGTTGGTGAATTAATGCGCCCGCGCAGAGATCAGTGGGAATTTTC 286
 Qy 302 TAGAACAATGCGAACAATTAATAATCAACAATAACAGAAATGCTAGGAATACGGCAC 361
 Db 287 TAGAACAATGCGAACAATTAATAATCAACAATAACAGAAATGCTAGGAATACGGCTC 346
 Qy 362 TTGCTGATTTACAGGTTTAGGATTTCTTTAGAGCTATCAACAGTCACTTTGAAGATT 421
 Db 347 TTGCTGATTTACAGGTTTAGGATTTCTTTAGAGCTATCAACAGTCACTTTGAAGATT 406
 Qy 422 GGCTAGAAAACCGTGATGATGCAAGAACGAGAAGTTCTTTATACCCAAATATATAGCCT 481
 Db 407 GGCTAGAAAACCGTGATGATGCAAGAACGAGAAGTTCTTTATACCCAAATATATAGCCT 466
 Qy 482 TAGAACHTGAATTTCTTAATGCGATGCGCTTTTCGCAATTAGAACAAGAGTTCCAT 541
 Db 467 TAGAACHTGAATTTCTTAATGCGATGCGCTTTTCGCAATTAGAACAAGAGTTCCAT 526
 Qy 542 TATTAATGTTATATGCTCAAGCTGCAAAATTTACACCTTATTAATTCAGAGATGCTCTC 601
 Db 527 TATTAATGTTATATGCTCAAGCTGCAAAATTTACACCTTATTAATTCAGAGATGCTCTC 586
 Qy 602 TTTTGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACGTTATTTATGAGCGCAAG 661
 Db 587 TTTTGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACGCTATTAATGAGCGCAAG 646
 Qy 662 TGGAAACAGAGATATTTCCGACTATTTGCGTAGAATGGTATAATACAGTCTAAATA 721
 Db 647 TGGAAACAGAGATATTTCCGACTATTTGCGTAGAATGGTATAATACAGTCTAAATA 706
 Qy 722 GCTTTGAGAGGACAAATGCGCAAGTTGGGTGCGTTATTAATCAATTCGTTAGAGATCTAA 781
 Db 707 GCTTTGAGAGGACAAATGCGCAAGTTGGGTGCGTTATTAATCAATTCGTTAGAGATCTAA 766
 Qy 782 CGTTAGGGGTATTAGACTAGTGGCACTATTTCCCAAGCTATGACATCGCACTTATCCAA 841
 Db 767 CGTTAGGGGTATTAGACTAGTGGCACTATTTCCCAAGCTATGACATCGCACTTATCCAA 826
 Qy 842 TAAATACAGTGTCTAGTTAACAAGGAAGTTTATACAGACGCAATTTGGAGCAACAGGG 901
 Db 827 TAAATACAGTGTCTAGTTAACAAGGAAGTTTATACAGACGCAATTTGGAGCAACAGGG 886
 Qy 902 TAAATATGCGAAGTATCAATTTGGTATTAATATATGCACTTCGTTTTCGCTATAGAGA 961
 Db 887 TAAATATGCGAAGTATCAATTTGGTATTAATATATGCACTTCGTTTTCGCTATAGAGG 946
 Qy 962 CTGCGGTATCCGAAGCCCGCATCTACTGATTTTCTAGAACAACTTACAAATTTTAGCA 1021
 Db 947 CTGCGGTATCCGAAGCCCGCATCTACTGATTTTCTAGAACAACTTACAAATTTTAGCG 1006
 Qy 1022 CTTTCATACGATGAGTGTCTACTAGGATATGACTTACTGCGGGGGGCAACAAATTCAT 1081
 Db 1007 CTTTCATACGATGAGTGTCTACTAGGATATGACTTACTTATGCGGGGGGCAACCGATTCAT 1066

Qy 1082 CTGCGCAATAGGAGCGGATTAATACTCTCAACGATGGGTCTACCAATCTCTATTA 1141
 Db 1067 CTGCGCAATAGGAGCGGATTAATACTCTCAACGATGGGTCTACCAATCTCTATTA 1126
 Qy 1142 ATCCTGTAGATTAATCATTTCTCTCGAGACGTATTTGGACTGAATCATATGCAAGG 1201
 Db 1127 ATCCTGTAAATTAATCATTTCTCGAGACGTATTTAGGACTGAATCATATGCAAGG 1186
 Qy 1202 TGCTTCTATGGGAAATTTACCTTTGAACCTTATCATGGTGTCCCTACTGTGTAGATTTAAT 1261
 Db 1187 TGCTTCTATGGGAAATTTACCTTTGAACCTTATCATGGTGTCCCTACTGTGTAGATTTAAT 1246
 Qy 1262 TTAGGAACCTCTAGATTAATTTTGAAGAGGTACTGTAACTATAGTCAACCTCTAGT 1321
 Db 1247 TTAGGAACCTCTAGATTAATTTCTGATAGAGTACCGCTAATATAGTCAACCTCTAGT 1306
 Qy 1322 CACTCGGCTTCAATTTAAAGATTCAGAACTGAAATTAACCAAGAAACAAACAGAACGAC 1381
 Db 1307 CACTCGGCTTCAATTTAAAGATTCAGAACTGAAATTAACCAAGAAACAAACAGAACGAC 1366
 Qy 1382 CAAATTAATGATCATATAGTATAGTATCTCACATAGGCTCANTTTCACAAATCTAGGG 1441
 Db 1367 CAAATTAATGATCATATAGTATAGTATCTCACATAGGCTAATTTTACAAATCTAGGG 1426
 Qy 1442 TGCAATGACAGTATATTTCTTGACGACACCGTATAGTGCAGATCGTACAAATCTAGTT 1501
 Db 1427 TGAATGTAACGATATATTTCTTGACGACACCGTATAGTGCNACTCTTACAAATCTAGT 1486
 Qy 1502 CAGATGATTAACAAATAACCAATTTGGTAAATCAATTTCAACCTTAATTTAGGTACCTCTG 1561
 Db 1487 CAGAGAAATTAATCAATTAATTTAGTGAAGAGATTTAGAGTTTGGGGGGGACCTCTG 1546
 Qy 1562 TACTGATGCGCCAGAGATTTACAGAGGGGATATAATTCGAACTAAAGTTTAAATGGTAGTG 1621
 Db 1547 TCAATCAGGACAGAGATTTACAGAGGGGATATAATTCGAACTAAAGTTTAAATGGTAGTG 1606
 Qy 1622 TACTAATGATGGCTTTAATTTTAAATAACATCAATTTACAGCGGTATTCGCGTACAGAGTTC 1681
 Db 1607 TTTGATCTTACAGTCAATTAATTTTACCAATTTACCAAGATACCGTTTAAAGATTTC 1666
 Qy 1682 GTTATGCTGCTTCT-----CAACAATGGTCTCTGAGGGTAACTG 1720
 Db 1667 GTTATGCTTCCAGTAGGGATGCAAGGATTAATTAACAGAGGGGATCCACAGGAG 1726
 Qy 1721 TCGAGGAGTACTACTTTTGTCAAGGATTCCTAGTACTATGAGTGAAGTGAAGTCTT 1780
 Db 1727 TGGAGCCCAAGTTAGTGAATATGCTCTTCAGAAACTATGGAATAGGGGAGACT 1786
 Qy 1781 TGACATCTCAATCAATTTAGATTTGCAAGATTTCTGTAGGTATTTAGTGCACTCTGGCAGTC 1840
 Db 1787 TAACATCTAGAACATTTAGATATACCGATTTTAGT---AATCTTTTCTTTTAGAGCTA 1843
 Qy 1841 AATCTGCTGGAATTAAGTAAAGTAA-----TAATGCAAGTAGAC 1879
 Db 1844 ATCCAGATATAATTTGGGATAAGTGAACCACTCTATTTGGTGCAGGTTCTTATTAGTAGG 1903
 Qy 1880 ABAAGTTTCACTTTGATAAATTAATTTCAATTTCTCAATTTCTGCAACCTTCGAAAGCAGAT 1939
 Db 1904 GTGAATTTTATATAGATAAATTTGAATTTATTTCTGAGATGCAACATTTGAGAGCAGAT 1963
 Qy 1940 ACGATTTAGAAAGGGCGCAAGAGCGGTGAATGCTCTGTTTACTTAATPACGAATCCAGAA 1999
 Db 1964 CTGATTTAGAAAGAGCAGCAAAAGCGGTGAATGCCCTGTTTACTTCTTCCAATCAATTCG 2023
 Qy 2000 GATTGAAAACAGATGTCAGAGATATCATATTTGATCAAGTATCCAATTTAGTGGCGTGT 2059
 Db 2024 GGTTAATAAACCAGTATGACGAGTATCATATTTGATCAAGTATCCAATTTAGTGGATGTT 2083
 Qy 2060 TATCGGATGAATTTCTGCTTAGATAAAGAGAAATTTACTTTGAGAAAGTCAAAATATCGGA 2119
 Db 2084 TATCAGATGAATTTTGTCTGATAAAGAGAGATTTGTCGAGAAAGTCAAAATATCGGA 2143
 Qy 2120 AACGACTCAGTGTAGAAAGAAACTTACTCCAAGATCCAACTTCAACATCCATCAATNAGC 2179

Db	2144		AGCGACTCAGTGAATGAGCGGAATTTACTTCAAGATCCAACTTCAGAGGATCAATAGAC	2203
Qy	2180		AACAGACTTCATATCTACTAATGAGCAATCGAAATTTACATCTATCCATGAAACAATCTG	2239
Db	2204		AACCAGAC-----	2211
Qy	2240		AACATGATGCTGGGAGTGCAGAACATTTACAATCCAGGAGGAATGACGTATTTAAAG	2299
Db	2212		--CGTGGCTGGAGAGGAAGTACAGATATTACCATCCAAAGGAGAGATGACGTATTTCAAAG	2269
Qy	2300		AGAATTACGTACACTACCGGGGACTTTTAATGAGTGTATCCGACGTATTTATATCAAA	2359
Db	2270		AGAATTACGTACACTACCGGGTACCGTGTGATGAGTGTATCCAAAGTATTTATATCAGA	2329
Qy	2360		AAATAGGAGAGTCCGAATTTAAAGCTTATCTCCCTACCAATTAAGAGGGTATATTGAAG	2419
Db	2330		AAATAGATGAGTCCGAATTTAAAGCTTATACCCGTATTGAATTAAGAGGGTATATCGAAG	2389
Qy	2420		ATAGTCAAGATTTAGAGATATATTGATTCGTTATATGCGAAACATGAACAATTTGGATG	2479
Db	2390		ATAGTCAAGACTTAGAAATCTATTGATCCGTTACAATGCAAAACACGAAATAGTAAAT	2449
Qy	2480		TTCCAGGTACCGAGTCCGATGCGCGCTTTTCAGTTGAAAGCCCAATCGGAGGTGCGGAG	2539
Db	2450		TGCCAGGACCGGTTCTTATGCGCGCTTTTCAGCCCAAGTCCCAATCGGAAAGTGTGGAG	2509
Qy	2540		AACCGAATTCGATGCGCACCAATTTTGAATGGAATCCTGATCTAGATTTGTTCTTCGACAG	2599
Db	2510		AACCGAATTCGATGCGCACCAATTTTGAATGGAATCCTGATCTAGATTTGTTCTTCGACAG	2569
Qy	2600		ATGAGAGAAATATGTCGATCATTTCCCATCAATTTCTCTTTGGATTTGATTTGGATGCA	2659
Db	2570		ACGGGAGAAATATGTCATCATTTCCCATCAATTTCCATTTGATTTGATTTGATTTGATTTG	2629
Qy	2660		CAGACTTCCATGAGAACTTAGCGTGTGGGTGATTTCAAGTTAAAGCCAGGAAGTTC	2719
Db	2630		CAGACTTAAATGAGGACTTAGGTGTATGGGTGATTTCAAGTTAAAGCCAGGAAGTTC	2689
Qy	2720		ATGCAAGACTAGGGAATCTGGAAATTTTAAAGAGAAACCAATTTATAGGAGAACCACTGT	2779
Db	2690		ATGCAAGACTAGGGAATCTAGATTTCTGAGAGAGAAACCAATTTATAGGGAAGCACTAG	2749
Qy	2780		CTCGTGTGAAGAGAGCAGAGAAAAATGGAGAGACAAACGTGAAAACTACAAATGGAAA	2839
Db	2750		CTCGTGTGAAGAGAGCAGAGAAAAATGGAGAGACAAACGTGAAAACTACAAATGGAAA	2809
Qy	2840		CAAAACAGATATATACAGAGCAAAAGAGCTGTGGATGCTTTATTTGTAGATTTCTCAAT	2899
Db	2810		CAAAATATTTGTTATAAAGAGCAAAAGAACTCTGATAGATGCTTTATTTGTAACTCTCAAT	2869
Qy	2900		ATAATAGATTACAAGCGGATACAAACATTTGGCATGATTCATGCGGCAGATAAACTTGTTC	2959
Db	2870		ATGATAGATTACAAGTGGATAGCAACATTCGGATGATTCATGCGGCAGATAAACCGGTTTC	2929
Qy	2960		ATCGAATTCGAGAGGCTTATCTGTGAGAAATATCTGTTATCCCGGGGTGTAATCCGAAA	3019
Db	2930		ATGAATCCGGGAAGCGTATCTGCCAGATTTGTTCTGATTTCCAGGTGTCAATGCGGCCA	2989
Qy	3020		TTTTTGAAGAAATTAGAGGTCGCAATTTACTGCAATCTCCCTATACGATCGCGAAATG	3079
Db	2990		TTTTTCGAAGAAATTAGAGGAGCGTATTTTTCAGCGTATTTCTTATATGATCGCGAAATG	3049
Qy	3080		TCGTTTAAATGTTGATTTTAAATATGATTTAGCATCTGGAATGTAAGAGGCGATGTAG	3139
Db	3050		TCATTTAAATGTTGCGATTTTCAATATGCGTATTTATGCTGGAACGTGAAAGGTCTATG	3109
Qy	3140		ATGTACAA---CAGAGCCATCACCGTTCGTCTTGTGTTATCCCAAGATGGGAAGCAGAAG	3196
Db	3110		ATGTAGAGAGCAAAACAAACACCGTTCGCTCTTGTGTTATCCCAAGATGGGAGCAGAG	3169
Qy	3197		TGTCACAGCAATTTGCGGTGTCTCGGGGCGTGTATATCTCCGTGTACAGCGTACA	3256

Db	3170		TGTCACAAAGAGTTTCGTGTCTGTCCAGGTCTGTGCTATATCTCTTCGTGTACAGCATATA	3229
Qy	3257		AAGAGGATATGAGAGGGTGTGTAAAGATCCATGAAATCGAGAACAAATACAGACGAAC	3316
Db	3230		AAGAGGATATGAGAGGGTGTGTAAAGATCCATGAAATCGAGAACAAATACAGACGAAC	3289
Qy	3317		TAAATTTTAAATACTGTGAAGAGAGGAGTGTATCCAAACGATACAGGAAACGTGTAATG	3376
Db	3290		TGAATTCAGCACTGTGTAGAGAGGAGTATATCCAAACACACAGTAAACGTGTAATA	3349
Qy	3377		ATTATATCTGCACACCAAGGTACAGCATATGT-----AATTCCTGTAATGCTCGATATG	3430
Db	3350		ATTATATCTGGGACTCAAGAGAAATATGAGGTCAGTACACTTCTCGTAAATCAAGSATATG	3409
Qy	3431		AGGATGCATATGAAGTTGATATCTACAGCATCTGTTAATTAATTAACCAACCGACTTATGAAGAG	3490
Db	3410		ACGAAGCTTATGGTAAATACCTTCCGTACCGACTGATTAGCTTCAAGTCTATGAAGAAA	3469
Qy	3491		AAACGTATACAGATGTACGAAGAGATATCATTTGTGTAATATGACAGAGGGTATGTGAATT	3550
Db	3470		AATCGTATACAGATGTGGAAGAGAGAAATCTTGTGTAATCTAAACAGAGGCTATGGGATT	3529
Qy	3551		ATCCACCACTACAGCTGCTTATATATGACAAAAGAAATTAAGAAATCTCCAGAAAACCGATA	3610
Db	3530		ACACCACTACCGGCTGTTTATGTAACAAGGATTTAGAGTACTTCCAGAGACCGATA	3589
Qy	3611		AGGTATCGATTTGAGATTTGAGAAAACGGAAGGAGTTTATTTGTAGACAGCGTGAATTTAC	3670
Db	3590		AGGTATCGATTTGAGATCGGAGAAAACAGAGGAAACATTCATCGTGGATAGCGTGAATTTAC	3649
Qy	3671		TCCTTATGGAGGAA	3684
Db	3650		TCCTTATGGAGGAA	3663

Search completed: April 20, 2005, 20:18:03
 Job time : 1771 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2005, 19:20:38 ; Search time 611 Seconds
(without alignments)
9873.899 Million cell updates/sec

Title: US-10-614-524-1

Perfect score: 3687

Sequence: 1 ttgacttcaataggagaaaaa.....tactctctatggagggaatag 3687

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3666.2	99.4	4173	4	US-09-661-322A-37
2	3266.6	88.6	3684	1	US-08-488-170-7
3	3266.6	88.6	3684	1	US-08-961-803-5
4	3260.2	88.4	3684	4	US-09-661-322A-62
5	3133.6	85.0	4074	1	US-08-377-690-1
6	2687.2	72.9	3934	1	US-08-100-709-3
7	2687.2	72.9	3934	1	US-08-176-865-3
8	2687.2	72.9	3934	1	US-08-474-038-3
9	2687.2	72.9	3934	2	US-08-779-046-3
10	2687.2	72.9	3934	2	US-08-881-340-3
11	1793.4	48.6	3522	1	US-08-040-751-4
12	1793.4	48.6	3522	1	US-08-291-368-1
13	1793.4	48.6	3522	2	US-08-962-190-1
14	1793.4	48.6	3522	5	PCT-US95-10310-1
15	1791.8	48.6	3522	6	5164180-3
16	1791.8	48.6	3522	6	5164180-3
17	1662	45.1	4106	1	US-08-434-823-1
18	1662	45.1	4106	1	US-08-457-366-1
19	1581.6	42.9	3558	3	US-08-178-252-22
20	1581.6	42.9	3558	4	US-08-826-660-22
21	1501	40.7	3766	1	US-08-032-364-1
22	1495.6	40.6	3624	1	US-07-951-715A-6
23	1495.6	40.6	3624	2	US-08-459-448A-6
24	1495.6	40.6	3624	3	US-08-459-595A-6
25	1495.6	40.6	3624	3	US-08-459-504B-6
26	1495.6	40.6	3624	3	US-08-459-444-6
27	1495.6	40.6	3624	3	US-08-053-549-7

28	1495.6	40.6	3624	3	US-09-547-422-6	Sequence 6, Appli
29	1495.6	40.6	3624	4	US-09-988-462-6	Sequence 6, Appli
30	1438.6	39.0	8854	3	US-09-053-549-1	Sequence 1, Appli
31	1432.4	38.9	3713	1	US-08-100-709-1	Sequence 1, Appli
32	1432.4	38.9	3713	1	US-08-176-865-1	Sequence 1, Appli
33	1432.4	38.9	3713	1	US-08-474-038-1	Sequence 1, Appli
34	1432.4	38.9	3713	2	US-08-779-046-1	Sequence 1, Appli
35	1432.4	38.9	3713	2	US-08-881-340-1	Sequence 1, Appli
36	1411.4	38.3	4020	5	PCT-US91-02560-1	Sequence 1, Appli
37	1408.4	38.2	3522	1	US-07-828-788A-9	Sequence 9, Appli
38	1408.4	38.2	3522	1	US-08-349-867-24	Sequence 24, Appli
39	1408.4	38.2	3522	1	US-08-239-476-24	Sequence 24, Appli
40	1408.4	38.2	3522	1	US-08-356-034-7	Sequence 7, Appli
41	1408.4	38.2	3522	1	US-08-598-305A-24	Sequence 24, Appli
42	1408.4	38.2	3522	2	US-08-639-923A-24	Sequence 24, Appli
43	1408.4	38.2	3522	3	US-08-933-891-7	Sequence 7, Appli
44	1408.4	38.2	3522	3	US-09-178-252-5	Sequence 5, Appli
45	1408.4	38.2	3522	4	US-09-521-344-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-661-322A-37
; Sequence 37, Application US/09661322A
; Patent No. 6593293
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Rupar, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin Composi
; TITLE OF INVENTION: and Methods of Use
; FILE REFERENCE: MECO201
; CURRENT APPLICATION NUMBER: US/09/661.322A
; CURRENT FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 4173
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3687)
US-09-661-322A-37

Query Match 99.4%; Score 3666.2; DB 4; Length 4173;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 3674; Conservative 0; Mismatches 13; Indels. 0; Gaps 0;

Qy	1	TTGACTTCAATAGGAAATGAGAAATTAATTAATGCTTTTATCGATTCAGCTGTA	60
Db	1	TTGACTTCAATAGGAAATGAGAAATTAATTAATGCTTTTATCGATTCAGCTGTA	60
Qy	61	TCGAATCATTTCCACACAAATGGATCTATCACCAGATGCTCGTATTGAGGATCTTTTGT	120
Db	61	TCGAATCATTTCCACACAAATGGATCTATCACCAGATGCTCGTATTGAGGATCTTTTGT	120
Qy	121	ATAGCCGAGGGGAATATATCAATCCACTTGTATTAGCGATCAACAGTCCAAACGGTATT	180
Db	121	ATAGCCGAGGGGAATATATCAATCCACTTGTATTAGCGATCAACAGTCCAAACGGTATT	180
Qy	181	AACATAGCTGTAGTAATAGGTATTAGCGTACGGTTGCTGACCAATAGCTAGT	240
Db	181	AACATAGCTGTAGTAATAGGTATTAGCGTACGGTTGCTGACCAATAGCTAGT	240
Qy	241	TTTTATAGTTTTCTTTGTTGGTGAATATATGCCCCCGCGCAGAGATCAGTGGGAAATTTTC	300
Db	241	TTTTATAGTTTTCTTTGTTGGTGAATATATGCCCCCGCGCAGAGATCAGTGGGAAATTTTC	300

QY 301 CTAGAACATGTCGAACAACTTATAAATCAACAAATAACAGAAAAATGCTAGGAATACGGCA 360
 DB |||||
 QY 301 CTAGAACATGTCGAACAACTTATAAATCAACAAATAACAGAAAAATGCTAGGAATACGGCA 360
 DB |||||
 QY 361 CTTGCTCGATTACAAGGTTTAGGAGATTCCTTTTAGAGCCTATCAACAGTCACTTTGAAGAT 420
 DB |||||
 QY 361 CTTGCTCGATTACAAGGTTTAGGAGATTCCTTTTAGAGCCTATCAACAGTCACTTTGAAGAT 420
 DB |||||
 QY 421 TGGCTAGAAAACCGTGATGTCGAAGAACGAGAGTGTCTTTTATATCCCAATATATAGCC 480
 DB |||||
 QY 421 TGGCTAGAAAACCGTGATGTCGAAGAACGAGAGTGTCTTTTATATCCCAATATATAGCC 480
 DB |||||
 QY 481 TTAGAACTTTGATTTCTTAATGCGATCGCGCTTTTTCGCAATAGAAACCAAGAGTTCCA 540
 DB |||||
 QY 481 TTAGAACTTTGATTTCTTAATGCGATCGCGCTTTTTCGCAATAGAAACCAAGAGTTCCA 540
 DB |||||
 QY 541 TTTAATATGTTATGTCCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGSCCTCT 600
 DB |||||
 QY 541 TTTAATATGTTATGTCCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGSCCTCT 600
 DB |||||
 QY 601 CTTTTTGTAGTGAATTTGGCTTTACATCGCAGGAAATTCACGTTATTATGAGCGCAA 660
 DB |||||
 QY 601 CTTTTTGTAGTGAATTTGGCTTTACATCGCAGGAAATTCACGTTATTATGAGCGCAA 660
 DB |||||
 QY 661 GTGAAACAAACGAGAGATTTATTCGACTATTGCGTAGAATGGTATAATACAGGTCTAAAT 720
 DB |||||
 QY 661 GTGAAACAAACGAGAGATTTATTCGACTATTGCGTAGAATGGTATAATACAGGTCTAAAT 720
 DB |||||
 QY 721 AGCTTGAGAGGGA CAATATGCGCAAGTTGGTGCGTTTAAATCAATTCCTGAGAGATCTA 780
 DB |||||
 QY 721 AGCTTGAGAGGGA CAATATGCGCAAGTTGGTGCGTTTAAATCAATTCCTGAGAGATCTA 780
 DB |||||
 QY 781 ACCTTGAGGTTATGATCTAGTGGCACTATTCGAAGCTATGACACTCGCACTTATCCA 840
 DB |||||
 QY 781 ACCTTGAGGTTATGATCTAGTGGCACTATTCGAAGCTATGACACTCGCACTTATCCA 840
 DB |||||
 QY 841 ATAAATACGAGTGTCTAGTTAAACAGGGAAGTTTATACAGACGCAATGGAGCAACAGGG 900
 DB |||||
 QY 841 ATAAATACGAGTGTCTAGTTAAACAGGGAAGTTTATACAGACGCAATGGAGCAACAGGG 900
 DB |||||
 QY 901 GTAATATGCGAAGTATGAATGGTATAATTAATGACCTTCGTTTTCCGCTATATAGAG 960
 DB |||||
 QY 901 GTAATATGCGAAGTATGAATGGTATAATTAATGACCTTCGTTTTCCGCTATATAGAG 960
 DB |||||
 QY 961 ACTCGGTTATCCGAAGCCGCATCTACTTGATTTCTAGAACCACTTGAATTTTATGAG 1020
 DB |||||
 QY 961 ACTCGGTTATCCGAAGCCGCATCTACTTGATTTCTAGAACCACTTGAATTTTATGAG 1020
 DB |||||
 QY 1021 ACTTCATCAGATGAGTGTCTAGGCATATGACTTACTGGCGGGGCGCACAAATTCAA 1080
 DB |||||
 QY 1021 ACTTCATCAGATGAGTGTCTAGGCATATGACTTACTGGCGGGGCGCACAAATTCAA 1080
 DB |||||
 QY 1081 TCTCGGCAATAGAGCGGANTTAAATACCTCAACGATGGGTCTACCAATCTTCTATT 1140
 DB |||||
 QY 1081 TCTCGGCAATAGAGCGGANTTAAATACCTCAACGATGGGTCTACCAATCTTCTATT 1140
 DB |||||
 QY 1141 AATCCTCTAGATTTATCATTTCTCTCGAGACCTATTTGGAAGTATATGACTGAATTCAGGA 1200
 DB |||||
 QY 1141 AATCCTCTAGATTTATCATTTCTCTCGAGACCTATTTGGAAGTATATGACTGAATTCAGGA 1200
 DB |||||
 QY 1201 GTGCTTCTATGGGAAATTTACTTTGAACCTTATTCATGGTGCTTCTAGTTAGATTTAAT 1260
 DB |||||
 QY 1201 GTGCTTCTATGGGAAATTTACTTTGAACCTTATTCATGGTGCTTCTAGTTAGATTTAAT 1260
 DB |||||
 QY 1261 TTTAGGAACCCCTCAGATATCTTTTGAAGAGGTAATCTGCTAATATAGTCAACCCCTATGAG 1320
 DB |||||
 QY 1261 TTTAGGAACCCCTCAGATATCTTTTGAAGAGGTAATCTGCTAATATAGTCAACCCCTATGAG 1320
 DB |||||
 QY 1321 TCACCTGGCTTCAATTTAAAGATTCAGAACTGAATTTACCACAGAAACCAACAGACGA 1380
 DB |||||
 QY 1321 TCACCTGGCTTCAATTTAAAGATTCAGAACTGAATTTACCACAGAAACCAACAGACGA 1380
 DB |||||
 QY 1381 CCAAAATTAATCATATAGTATCATAGTTATCTCACATAGGGCTCAATTCACAAATCTPAGG 1440
 DB |||||

DB 1381 CCAAAATTAATCATATAGTATCATAGTTATCTCACATAGGGCTCATTTCCAAATCTAGG 1440
 QY |||||
 DB 1441 GTGCATGTAACAGTATATCTTGGACGACCGTAGTGCAGATCGTACAAATACCAATAGT 1500
 QY |||||
 DB 1441 GTGCATGTAACAGTATATCTTGGACGACCGTAGTGCAGATCGTACAAATACCAATAGT 1500
 QY |||||
 QY 1501 TCAGATAGCATAACAAATACCAATACCAATTCGTTAAATCATTTCAACCTTAATTCAGGTACCTCT 1560
 DB |||||
 QY 1501 TCAGATAGCATAACAAATACCAATTCGTTAAATCATTTCAACCTTAATTCAGGTACCTCT 1560
 DB |||||
 QY 1561 GTAGTCAAGTGGCCAGGATTTACAGGAGGGAATATAATCCGAACCTAACGTTAATGGTAGT 1620
 DB |||||
 QY 1561 GTAGTCAAGTGGCCAGGATTTACAGGAGGGAATATAATCCGAACCTAACGTTAATGGTAGT 1620
 DB |||||
 QY 1621 GTACTAAGTATGGTCTTAATTTAATAATACATATTAACAGCGGTATCCGCTGAGAGTT 1680
 DB |||||
 QY 1621 GTACTAAGTATGGTCTTAATTTAATAATACATATTAACAGCGGTATCCGCTGAGAGTT 1680
 DB |||||
 QY 1681 CGTTATGCTGCTTCTCAACAAATGCTCCTGAGGTAACCTGTCGAGGGAAGTACTACTTTT 1740
 DB |||||
 QY 1681 CGTTATGCTGCTTCTCAACAAATGCTCCTGAGGTAACCTGTCGAGGGAAGTACTACTTTT 1740
 DB |||||
 QY 1741 GATCAAGGATTCCTCTAGTACTATGAGTGCATAATGAGTCTTTTGACATCTCAATCATTTAGA 1800
 DB |||||
 QY 1741 GATCAAGGATTCCTCTAGTACTATGAGTGCATAATGAGTCTTTTGACATCTCAATCATTTAGA 1800
 DB |||||
 QY 1801 TTTGCGAATTTCTGTAGGTAATGAGTGCATCTGGCAGTCAAACTGCTGGAATAAGTATA 1860
 DB |||||
 QY 1801 TTTGCGAATTTCTGTAGGTAATGAGTGCATCTGGCAGTCAAACTGCTGGAATAAGTATA 1860
 DB |||||
 QY 1861 AGTAATAATSCAGTAGACAAACGTTTCACTTTGATAAAATTCGAATTCATTCCAAATFAC 1920
 DB |||||
 QY 1861 AGTAATAATSCAGTAGACAAACGTTTCACTTTGATAAAATTCGAATTCATTCCAAATFAC 1920
 DB |||||
 QY 1921 GCAACCTTCTGAAACGAGATA CGATTTAGAAAGGCGCAAGAGCGGTGAATTCCTGTTTT 1980
 DB |||||
 QY 1921 GCAACCTTCTGAAACGAGATA CGATTTAGAAAGGCGCAAGAGCGGTGAATTCCTGTTTT 1980
 DB |||||
 QY 1981 ACTAATACGAATCCAGAGATTTGAAACAGATGTGACAGATTTATCAATTTGATCAAGTA 2040
 DB |||||
 QY 1981 ACTAATACGAATCCAGAGATTTGAAACAGATGTGACAGATTTATCAATTTGATCAAGTA 2040
 DB |||||
 QY 2041 TCCAAATTTAGTGGCTGTTTTATCGATGAATTCCTGTAGATGAAAGAGAGATTTACTT 2100
 DB |||||
 QY 2041 TCCAAATTTAGTGGCTGTTTTATCGATGAATTCCTGTAGATGAAAGAGAGATTTACTT 2100
 DB |||||
 QY 2101 GAGAAAGTGAATATGCAAAACGACTCAGTGTGAAAGAACTTACTCCAAAGATCCAAAC 2160
 DB |||||
 QY 2101 GAGAAAGTGAATATGCAAAACGACTCAGTGTGAAAGAACTTACTCCAAAGATCCAAAC 2160
 DB |||||
 QY 2161 TCCATATCCATCAATTAAGCAACAGATTTCAATCTACTAATGAGCAATCGAATTTTACA 2220
 DB |||||
 QY 2161 TCCATATCCATCAATTAAGCAACAGATTTCAATCTACTAATGAGCAATCGAATTTTACA 2220
 DB |||||
 QY 2221 TCTATCCATGAACAACTGAAACGATGATGTTGGGGAAGTGAAGAACTTCAATCCAGGAA 2280
 DB |||||
 QY 2221 TCTATCCATGAACAACTGAAACGATGATGTTGGGGAAGTGAAGAACTTCAATCCAGGAA 2280
 DB |||||
 QY 2281 GGAATGACGTATTTAAAGAGAAATACGTCACACTACCGGGGACTTTTAAATGAGTGTAT 2340
 DB |||||
 QY 2281 GGAATGACGTATTTAAAGAGAAATACGTCACACTACCGGGGACTTTTAAATGAGTGTAT 2340
 DB |||||
 QY 2341 CCGAGCTATTATATCAAAAAATAGAGAGTCCGAATTTAAAGCTTATATCTCGCTACCAA 2400
 DB |||||
 QY 2341 CCGAGCTATTATATCAAAAAATAGAGAGTCCGAATTTAAAGCTTATATCTCGCTACCAA 2400
 DB |||||
 QY 2401 TTAAGAGGGTATATTCGAAGTAGTCAAGATTTTAGAGATATATTTGATTCGTATATATGCG 2460
 DB |||||
 QY 2401 TTAAGAGGGTATATTCGAAGTAGTCAAGATTTTAGAGATATATTTGATTCGTATATATGCG 2460
 DB |||||
 QY 2461 AAAACATGAACATTTGATGTTCCAGGTACCGAGTCCGATGCGCGCTTTCAGTTGAAAGC 2520
 DB |||||

Qy	2275	CAGGAAGAAATGACGTAATTTAAAGAGAAATTAACGTCACTACCGGGACCTTTTAATGAG	2334
Db	2272		
		CAGGAAGAAATGACGTAATTTAAAGAGAAATTAACGTCACTACCGGGTACTTTTAATGAG	2331
Qy	2335	TGTTATCCGACGTATTTATATCAAAAAATAGGACAGTCGGAATTTAAAGCTTATACCTCGC	2394
Db	2332		
		TGTTATCCGACGTATTTATATCAAAAAATAGGGAGCGGAATTTAAAGCTTATACCTCGC	2391
Qy	2395	TACCAATTAAGAGGGTATATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT	2454
Db	2392		
		TACCAATTAAGTGGCTATATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTAT	2451
Qy	2455	AATCGGAAACATGAAAAATTTGGATGTTTCCAGGTTACCGAGTCGGTATGCGCGCTTTCAGTT	2514
Db	2452		
		AATCGGAAACATGAAAAATTTGGATGTTTCCAGGTTACCGAGTCGGTATGCGCGCTTTCAGTT	2511
Qy	2515	GAAGCCCAATCGGAAGGTGCGGAGAACCGAATCGATCGCACACACATTTTGAATGGAAT	2574
Db	2512		
		GAAGCCCAATCGGAAGGTGCGGAGAACCGAATCGATCGCACACATTTTGAATGGAAT	2571
Qy	2575	CCTGATCTAGATGTTTCTCGCAGAGATGGAGAAAAATGTGCGCATCATTTCCCATCATTTTC	2634
Db	2572		
		CCTGATCTAGATGTTTCTCGCAGAGATGGAGAAAAATGTGCGCATCATTTCCCATCATTTTC	2631
Qy	2635	TCITTTGATATTTGATATTTGATGTCACAGACTTGTGATGAGAAATCTAGCGGTGTGGGTGGA	2694
Db	2632		
		TCITTTGATATTTGATATTTGATGTCATAGACTTGTGATGAGAACCTTAGCGGTGTGGGTGGA	2691
Qy	2695	TTCAAGATTAAAGCGCAGGAGGTCATCGAAGACTAGGGAATCTGGAAATTTATTTAGAGAG	2754
Db	2692		
		TTCAAGATTAAAGCGCAGGAGGTCATCGAAGACTAGGGAATCTGGAAATTTATTTAGAGAG	2751
Qy	2755	AAAACTATTATTAGGAAGCACTGTCGTGTGAAGAGAGCAGAGAAAAAATGGAGAGAC	2814
Db	2752		
		AAACCATATTAGGAAGCACTGTCGTGTGAAGAGAGCAGAGAAAAAATGGAGAGAC	2811
Qy	2815	AAACGTGAAAAAATCAATTTGGAAAAAACAACGAGTATATACAGAGGCAAAAGAGCTGTG	2874
Db	2812		
		AAACGTGAAAAAATCAATTTGGAAAAAACAACGAGTATATACAGAGGCAAAAGAGCTGTG	2871
Qy	2875	GATGCTTTATTTGATAGATTTCTCAATATATAGATTACAAGCGGATACAAACATTTGGCATG	2934
Db	2872		
		GATGCTTTATTTGATAGATTTCTCAATATATAGATTACAAGCGGATACAAACATTTGGCATG	2931
Qy	2935	ATTATCGCGCAGATAAATCTGTTTCATCGAAATTCGAGAGCTTATCTGTGTCAGAAATATCT	2994
Db	2932		
		ATTATCGCGCAGATAAATCTGTTTCATCGAAATTCGAGAGCTTATCTGTGTCAGAAATATCT	2991
Qy	2995	GTTATCCGGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAAGTGCATATTACACTGCA	3054
Db	2992		
		GTTATCCCGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAAGTGCATATTACACTGCA	3051
Qy	3055	ATCTCCCTATACGATGCGAGAAATGTCGTTAAAAATGGTGATTTTAATTAATGATGATAGCA	3114
Db	3052		
		ATCTCCCTATACGATGCGAGAAATGTCGTTAAAAATGGTGATTTTAATTAATGATGATAGCA	3111
Qy	3115	TGCTGGAAATGTAAAGGGCATGTAGATGTACAACAGAGCCATCACCGTCTGTCTGCTTGT	3174
Db	3112		
		TGCTGGAAATGTAAAGGGCATGTAGATGTACAACAGAGCCATCACCGTCTGTCTGCTTGT	3171
Qy	3175	ATCCGAAATGGGAAGCAGAAAGTGTCAACAGCAGATTCGCGTCTGTCTCGGGGCGTGGCTAT	3234
Db	3172		
		ATCCGAAATGGGAAGCAGAAAGTGTCAACAGCAGATTCGCGTCTGTCTCGGGGCGTGGCTAT	3231
Qy	3235	ATCTCCGTGTCAACGGTCAAAAGAGGATATGGAGAGGGTGTGTGTAACGATCCATGAA	3294
Db	3232		
		ATCTCCGTGTCAACGGTCAAAAGAGGATATGGAGAGGGTGTGTGTAACGATCCATGAA	3291
Qy	3295	ATCGAGAACAAATACAGACGAACTAAAAATTTAAAACTGTGTAAGAGGAGGTGTATCCA	3354
Db	3292		
		ATCGAGAACAAATACAGACGAACTAAAAATTTAAAACTGTGTAAGAGGAGGTGTATCCA	3351

RESULTS

US-08-961-803-5
 : Sequence 5, Application US/08961803
 : Patent No. 6150589
 : GENERAL INFORMATION:
 : APPLICANT: Payne, Jewel
 : APPLICANT: Cummings, David A.
 : APPLICANT: Cannon, Raymond J.C.
 : APPLICANT: Narva, Kenneth E.
 : APPLICANT: Stelman, Steve
 : TITLE OF INVENTION: No. 6150589el Bacillus thuringiensis Isolate Denoted
 : TITLE OF INVENTION: B.t. PS158C2, Active Against Lepidopteran Pests, and Genes
 : TITLE OF INVENTION: Encoding Lepidopteran-Active Toxins
 : NUMBER OF SEQUENCES: 10
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Jay M. Sanders
 : STREET: 2421 N.W. 41st Street, Suite A-1
 : CITY: Gainesville
 : STATE: Florida
 : COUNTRY: USA
 : ZIP: 32606
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/961,803
 : FILING DATE: 31-OCT-1997
 : CLASSIFICATION: 800
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/069,902
 : FILING DATE: 01-JUNE-1993
 : CLASSIFICATION: 800
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/759,247
 : FILING DATE: 13-SEPT-1991
 : CLASSIFICATION: 800
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/448,170
 : FILING DATE: 23-MAY-1995
 : CLASSIFICATION: 800
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Sanders, Jay M.
 : REGISTRATION NUMBER: 39,355
 : REFERENCE/DOCKET NUMBER: W/S 102DCD1
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (352) 375-8100

```

; TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3684 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-961-803-5

```

Query Match 88.6%; Score 3266.6; DB 3; Length 3684;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 3449; Conservative 0; Mismatches 229; Indels 15;

Qy	1	TTGACTTCCAATAAGGAAAAATGAGAAATGAATAAATGCTTTATCAATCCAGCTGTA	60
Db	1	TTGACTTCCAATAAGGAAAAATGAGAAATGAATAAATGCTTTATCAATCCAGCTGTA	60
Qy	61	TCGAATCATTTCCACAAATGGATCTATCACAGATGCTCGTATTGAGGATTTCTTTGTGT	120
Db	61	TCGAATCATTTCCGACAAATGAACTCTATCAACCGATGCTCGTATTGAGATAGCTTTGT	120
Qy	121	ATAGCCGAGGGGAATAATATCAATCCACTGTTTAGCGCATCAACAGTCCAAACGGGTATT	180
Db	121	ATAGCCGAGGGGAACAATATCGATCTCATTTGTTAGCGCATCAACAGTCCAAACGGGTATT	180
Qy	181	AACATAGTGTGTAAGTAATACTAGGTGTAATTAGGCGTACCGTTTGCTGGACAAATAGCTAGT	240
Db	181	AACATAGTGTGTAAGTAATACTAGGTGTAATTAGGCGTACCGTTTGCTGGACAAATAGCTAGT	240
Qy	241	TTTTATAGTTTTCTTTGTTGGTGAATTTATGGCCCCCGCGCAGAGATCACTGGGAAATTTTC	300
Db	241	TTTTATAGTTTTCTTTGTTGGTGAATTTATGGCCCCCGCGCAGAGATCTCTGGGAAATTTTC	300
Qy	301	CTAGAACATGTCGAACAACTTTATAATCAACAAATAACAGAAATGCTTAGNATACGGCA	360
Db	301	CTAGAACATGTCGAACAACTTTATAAGACAAACAAGTAAACAGAAAAATACTAGGAGTACGGCT	360
Qy	361	CTTGCTCGATTACAAGGTTTAGGAGATTTCTTTTAGAGCTTATCAACAGTCACTTGAAGAT	420
Db	361	CTTGCTCGATTACAAGGTTTAGGAAATTTCTTTAGAGCTTATCAACAGTCACTTGAAGAT	420
Qy	421	TGGCTAGAAAAACCGTGATGATGCAAGACGAGAGTGTCTTTATACCAATATATAGCC	480
Db	421	TGGCTAGAAAAACCGTGATGATGCAAGACGAGAGTGTCTTTATACCAATATATAGCC	480
Qy	481	TTAGAACTTTGATTTTCTTAAATGCGATGCGCTTTTCGCAATTAGAAAAACAAGAGTTCCA	540
Db	481	TTAGAACTTTGATTTTCTTAAATGCGATGCGCTTTTCGCAATTAGAAAAACAAGAGTTCCA	540
Qy	541	TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGCCCT	600
Db	541	TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTGAGAGATGCCCT	600
Qy	601	CTTTTTGCTAGTGAATTTTGGGCTTACATCGCAGGAAATTTCAACGTTATTATGAGCGCCA	660
Db	601	CTTTTTGCTAGTGAATTTTGGGCTTACATCCCCAAGAAATTTCAACGTTATTATGAGCGCCA	660
Qy	661	GTGAAACAAACGAGAGATTTATTTCCGACTATTTCGCTAGAAATGGTATAATACAGGCTTAAT	720
Db	661	GTGAAACAAACGAGAGATTTCTGATTTATTTCGCAAGTGGTATAATACGGGTTTAAT	720
Qy	721	AGCTTAGAGGGAACAAATGCGCAAGTTGGGTGCGTTTATTAATCAATTCGTTAGAGATCTA	780
Db	721	AATTTTAGAGGGAACAAATGCTGAAATTTGGTTGGCATATAATCAATTCGTTAGAGACTTA	780
Qy	781	ACGTTTAGGGTATTAGATCTAGTGGCATTTTCCCAAGCTATGACACTCGCACTTATCCA	840
Db	781	ACGCTAGAGATTTAGATCTAGTGGCATTTTCCCAAGCTATGACACGGGTGTTTATCCA	840
Qy	841	ATAAATACGAGTGTCTAGTTTAAACAAGGGGAAGTTTATACAGACGCAATTTGGAGCAACAGG	900
Db	841	ATGAATACCGTGTCTCAATTTAAACAAGAGAAATTTTATACAGATTCGAATTTGGGAGAACAA	900

Qy	1975	CTGTTTACTAATACGAATCCAGAGAGTTTGAAAAACAGATGTGCAGAAATTATCATATTGAT	2034
Db	1972	CTGTTTACTAATACGAATCCAGAGAGTTTGAAAAACAGTGTGCAGAAATTATCATATTGAT	2031
Qy	2035	CAAGTATCCAAATTTAGTGGCGTGTTTATCGATGAAATTCCTGCTTAGATGAAAGAGAGAA	2094
Db	2032	GAAGTATCCAAATTTAGTGGCGTGTTTATCGATGAAATTCCTGCTTAGATGAAAGAGAGAA	2091
Qy	2095	TTACTTGAGAAAGTGAAATATGCGAAACGACTCAGTGTGATGAAAGAACTTTACTCCAAGAT	2154
Db	2092	TTACTTGAGAAAGTGAAATATGCGAAACGACTCAGTGTGATGAAAGAACTTTACTCCAAGAT	2151
Qy	2155	CCAAACTTCATCCATCAATTAAGCAACACGACTCTTCAATATCTACTAATGAGCAATCGAAT	2214
Db	2152	CCAAACTTCATCCATCAATTAAGCAACACGACTCTTCAATATCTACTAATGAGCAATCGAAT	2211
Qy	2215	TTCACTCTCATCCATGAACCAATCTGAACATCGATGTTGGGGAAGTGAGAACATTTACAATC	2274
Db	2212	TTCACTCTCATCCATGAACCAATCTGAACATCGATGTTGGGGAAGTGAGAACATTTACAATC	2271
Qy	2275	CAGGAAGGAAATGACGTATTTTAAAGAGAAATTAACGTCACTACCGGGGCACTTTTAAATGAG	2334
Db	2272	CAGGAAGGAAATGACGTATTTTAAAGAGAAATTAACGTCACTACCGGGGCACTTTTAAATGAG	2331
Qy	2335	TGTTATCCGACGTATTTATATCAAAAAATAGGAGAGTCGGAAATTAAGAAGCTTATATCTCGC	2394
Db	2332	TGTTATCCGACGTATTTATATCAAAAAATAGGAGAGTCGGAAATTTAAAGCTTATATCTCGC	2391
Qy	2395	TACCAATTAAGAGGATATTTGAAGTAGTCAAGATTTAGAGATATATTTGATTCGTTAT	2454
Db	2392	TACCAATTAAGAGGATATTTGAAGTAGTCAAGATTTAGAGATATATTTGATTCGTTAT	2451
Qy	2455	AATCGCAAAACATGAAACATTTGGATGTTTCCAGGTACCGAGTCCGCTATGGCGCGCTTTCAGTT	2514
Db	2452	AATCGCAAAACATGAAACATTTGGATGTTTCCAGGTACCGAGTCCGCTATGGCGCGCTTTCAGTT	2511
Qy	2515	GAAAGCCCAATCGGAAGGTCGGGAGAACCGAATCGATCGCGCAACCAATTTTGAATGGAAT	2574
Db	2512	GAAAGCCCAATCGGAAGGTCGGGAGAACCGAATCGATCGCGCAACCAATTTTGAATGGAAT	2571
Qy	2575	CCTGATCTAGATGTTCTCCTCAGAGATGGAGAAAAATGTCGGCATCAATCCCATCATTTTC	2634
Db	2572	CCTGATCTAGATGTTCTCCTCAGAGATGGAGAAAAATGTCGGCATCAATCCCATCATTTTC	2631
Qy	2635	TCCTTGATATTTGATATTTGGATGCACAGACTTGCATGAGAATCTAGCGCTGTGGGTGGTA	2694
Db	2632	TCCTTGATATTTGATATTTGGATGCATAGACTTGCATGAGAATCTAGCGCTGTGGGTGGTA	2691
Qy	2695	TTCAAGATTTAAGACGACGAGGAAGTTCATGCAAGACTAGGGAATCTGGAATTTTATGAAGAG	2754
Db	2692	TTCAAGATTTAAGACGACGAGGAAGTTCATGCAAGACTAGGGAATTTTATGAAGAG	2751
Qy	2755	AAACCATTTATAGGAGAGCACTGTCTCGTGTGAAGAGACGAGAAAAAATGGAGAGAC	2814
Db	2752	AAACCATTTATAGGAGAGCACTGTCTCGTGTGAAGAGAGCAGAAAAAATGGAGAGAC	2811
Qy	2815	AAACGTGAAAAAATCAAAATTTGAAAAACAAACGAGTATATACAGAGGCAAAAGAGCTGTG	2874
Db	2812	AAACGTGAAAAAATCAAAATTTGAAAAACAAACGAGTATATACAGAGGCAAAAGAGCTGTG	2871
Qy	2875	GATGCTTTATTTGATATCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATG	2934
Db	2872	GATGCTTTATTTGATATCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATG	2931
Qy	2935	ATTTCATCGGCAGATAAATCTGTTTCATCGAATTCGAGAGGCTTATCTGTACAGAAATTATCT	2994
Db	2932	ATTTCATCGGCAGATAAATCTGTTTCATCGAATTCGAGAGGCTTATCTGTACAGAAATTATCT	2991
Qy	2995	GTTATCCGGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAAAGTGCATTTATCACTGCA	3054
Db	2992	GTTATCCAGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAAAGTGCATTTATCACTGCA	3051
Qy	3055	ATCTCCCTTATACGATGCGAGAAATGTGCTTTAAAAAATGGTGAATTTTATAATGATTAGCA	3114

Db	3052	ATCTCCCTTACGATCGGAATATGCTGTAAAAATGGTGATTTTAATATAGGATTAGCA	3111
Qy	3115	TGCTGGAATGTAAGAGGCGATGTAGATGTAACAACAGAGCCATCACCGTTCTGTCCCTTGTT	3174
Db	3112	TGCTGGAATGTAAGAGGCGATGTAGATGTACAACAGAGCCATCACCGTTCTGTCCCTTGTT	3171
Qy	3175	ATCCCAATGGGAACGACGAGAGTGTCAACAGCAGTTTCGGTCTGTCCGGGGCGTGCTAT	3234
Db	3172	ATCCCAATGGGAACGAGAGTGTCAACAGCAGTTTCGGTCTGTCCGGGGCGTGCTAT	3231
Qy	3235	ATCCTCCGTGTACAGCGTACAAAGAGGGATATGGAGAGGGTTGTGTAAACGATCCCATGA	3294
Db	3232	ATCCTCCGTGTACAGCGTACAAAGAGGGATATGGAGAGGGTTGTGTAAACGATCCCATGA	3291
Qy	3295	ATCGAGAACAAATACAGACGAACTTAAATTTTAAAACTGTGAAGAAAGGAGTGTATCCA	3354
Db	3292	ATCGAGAACAAATACAGACGAACTTAAATTTTAAAACTGTGAAGAAAGGAGTGTATCCA	3351
Qy	3355	ACGGATACAGGAACGTTGTAATGATTACTGTGCACACCAAGGTACAGCAGTATGTAATTC	3414
Db	3352	ACGGATACAGGAACGTTGTAATGATTACTGTGCACACCAAGGTACAGCAGTATGTAATTC	3411
Qy	3415	CGTAATGCTGGATATGAGGATGCAATATGAAGTTGATACTACAGCATCTGTTAATTACAA	3474
Db	3412	CGTAATGCTGGATATGAGGATGCAATATGAAGTTGATACTACAGCATCTGTTAATTACAA	3471
Qy	3475	CCGACTTATCAAGAAAGACGTTATACAGATGTACGAAGAGATAATCATTTGTGAATATGAC	3534
Db	3472	CCGACTTATCAAGAAAGACGTTATACAGATGTACGAAGAGATAATCATTTGTGAATATGAC	3531
Qy	3535	AGAGGGTATGTGAATTATCCACCTACCAAGCTGGTTATATGACAAAGAAATTAGAATAC	3594
Db	3532	AGAGGGTATGTGAATTATCCACCAGTACCAGCTGGTTATATGACAAAGAAATTAGAATAC	3591
Qy	3595	TTCCCGAAACCGATAGGTATCGATTGAGATTGGAGAAAACGGAAGGAAAGTTTATTGTA	3654
Db	3592	TTCCCGAAACCGATAGGTATCGATTGAGATTGGAGAAAACGGAAGGAAAGTTTATTGTA	3651
Qy	3655	GACAGCGTGAATTACTCTCTTATGGAGGAATAG	3687
Db	3652	GACAGCGTGAATTACTCTCTTATGGAGGAATAG	3684

RESULT 4
US-09-661-322A-62
; Sequence 62. Application US/09661322A

```

: Patent No. 6593293
:
: GENERAL INFORMATION:
:
: APPLICANT: Baum, James A.
: APPLICANT: Chu, Chih-Rei
: APPLICANT: Donovan, William P.
: APPLICANT: Gilmer, Amy J.
: APPLICANT: Rupaar, Mark J.
:
: TITLE OF INVENTION: Lepidoptera
: TITLE OF INVENTION: and Method
: FILE REFERENCE: MECO201
: CURRENT APPLICATION NUMBER: US/
: CURRENT FILING DATE: 2000-09-1
: NUMBER OF SEQ ID NOS: 63
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 62
: LENGTH: 3684
: TYPE: DNA
: ORGANISM: Bacillus thuringiensis
US-09-661-322A-62

```

Query Match	88.4%;	Score 3260.2;	DB 4;	Length 3684;
Best Local Similarity	93.3%;	Pred. No. 0;		
Matches 3445:	Conservative	0;	Mismatches 233;	Indels 15;

QY 1 TTGACTTCAAAATAGGAAAAATGAGAATGAAATATATAATGCTTTATCGATTCCAGCTGTA 60

Db 1 TTGACTTCAAAATAGGAAAAATGAGAAATGAAATTAATAATGCTTTATCGAATCCAGCTGTA 60
 QY 61 TCGAATCAATCCACACAAATGGAATCTATCACCAGATGCTGTAATTTAGGATTTCTTGTGT 120
 Db 61 TCGAATCAATCCGACAAATGAAATCTATCAACCGATGCTGTAATGAGGATAGCTTGTGT 120
 QY 121 ATAGCCGAGGGGAATATATCAATCCACTTGTAGCGCATCAACAGATCCAAACGGGTATT 180
 Db 121 ATAGCCGAGGGGAACATATATCGATCCATTTGTAGCGCATCAACAGATCCAAACGGGTATT 180
 QY 181 AACATAGCTGTAGAAATCTATAGGTGTATTTAGCGGTACCGTTTGTGACAAATAGCTAGT 240
 Db 181 AACATAGCTGTAGAAATCTATAGGTGTATTTAGCGGTACCGTTTGTGACAAATAGCTAGT 240
 QY 241 TTTTATAGTTTCTTGTGTGTAATATGCGCGCGCGGAGAGATCAGTGGGAATTTTC 300
 Db 241 TTTTATAGTTTCTTGTGTGTAATATGCGCGCGCGGAGAGATCAGTGGGAATTTTC 300
 QY 301 CTAGAACATGTGGAACCAACTTATAAATCAACAAATTAACAGAAAATGCTAGGAATACGGCA 360
 Db 301 CTAGAACATGTGGAACCACTTATAAGACAAACAGTAACAGAAATCTAGGGATACGGCT 360
 QY 361 CTTGCTCGATTAAGAGTTTAGAGATTCCTTTAGAGCTATACAGTCACTTTGAAGAT 420
 Db 361 CTTGCTCGATTAAGAGTTTAGAGATTCCTTTAGAGCTATCAACAGTCACTTTGAAGAT 420
 QY 421 TGGCTAGAAAACCGTGATGTCGAAGACAGAGAGTGTCTTTATACCCCAATATATAGCC 480
 Db 421 TGGCTAGAAAACCGTGATGTCGAAGACAGAGAGTGTCTTTATACCCCAATATATAGCC 480
 QY 481 TTAGAACTTTGATTTTCTTAATGCGATGCGCTTTTCGCAATAGAAAGTTTCCA 540
 Db 481 TTAGAACTTTGATTTTCTTAATGCGATGCGCTTTTCGCAATAGAAAGTTTCCA 540
 QY 541 TTATTAATGATATGCTCAAGCTGCAAAATTTACACCTATTTATTTAGAGATGCTCT 600
 Db 541 TTATTAATGATATGCTCAAGCTGCAAAATTTACACCTATTTATTTAGAGATGCTCT 600
 QY 601 CTTTTCGTAGTGAATTTGGGCTTACATCGCAGAAATTCACCTTATTTATGAGCGCAA 660
 Db 601 CTTTTCGTAGTGAATTTGGGCTTACATCGCAGAAATTCACCTTATTTATGAGCGCAA 660
 QY 661 GTGGAACAAACAGAGATTTATTCGACTATTTGCGGTAGAAATGGTATAATACAGGTCTAAAT 720
 Db 661 GTGGAACAAACAGAGATTTATTCGATTTATGCGCAAGATGGTATAATACAGGTCTAAAT 720
 QY 721 AGCTTGAGAGGGAATAATGCGCAAGTTGGGTGCTTTATATCAATTTCCGTAGAGATCTA 780
 Db 721 AATTTGAGAGGGAATAATGCTGAAAGTTGGTGGCATATAATCAATTTCCGTAGAGACTTA 780
 QY 781 ACGTTAGGGTATTAGATCTAGTGGCACTATTTCCAGCTATGACACTCCGACTTATCCA 840
 Db 781 ACGTTAGGGTATTAGATCTAGTGGCACTATTTCCAGCTATGACACTCCGCTTTATCCA 840
 QY 841 ATAAATACGAGTGTCTAGTTAAACAGGGAAGTTTATACAGACGCAATTTGGAGCAACAGGG 900
 Db 841 ATGAATACCACTGTCTCAATTAACAGAGAAATTTATACAGATCCAATTTGGGAGAACAAAT 900
 QY 901 GTAAT-----ATGGCAAGTATGAATTTGGTATTAATTAATGACCTTGTTCGCT 954
 Db 901 GCACCTTCAGGATTTGCAAGTACGAATTTGGTATTAATTAATGACCACTGTTCGCTGCC 960
 QY 955 ATAGAGACTGGGTATTCGGAAGCCGCTATCTGTTGATTTTCGAGAACACTTACAAAT 1014
 Db 961 ATAGAGGCTGCGTTATTTAGGCTTCGCACTCTGATTTTCGAGAACACTTACAAAT 1020
 QY 1015 TTTAGCACTTCATCAGATGAGTGTCTACTAGGCAATATGACTTACTGGCGGGGCAACA 1074
 Db 1021 TTCAGCGTATTAAGTTCGATGAGTAATCTCAATATATGAATTTACTGGTGGGACATAGA 1080
 QY 1075 ATTCAACTTCGCGCAATAGGAGCGGATTAATACCTCAACGATGGTCTACCAATACT 1134
 Db 1081 CTTGAATCGCGAACAAATAAGGGGTCATTAAGTACCTGGACACACCGGAAATACCAATACT 1140

QY 1135 TCTATTAAATCCTGTAGATTTATCAATCTCTCTCGAGACGTATATTGCACTGAATCATAT 1194
 Db 1141 TCTATTAAATCCTGTAAATTTACAGTTTCAATCTCGAGACGTATTATAGAACAGAAATCATTT 1200
 QY 1195 GCAGGAGTGTCTTATGGGAAATTTACCTTGAACCTTATTCATGTTGCTCCCTACTGTTAGA 1254
 Db 1201 GCAGGATAAATAT-----ACTTCTAACTACTCTCTGTAATGGAGTACCTTGGGCTAGA 1254
 QY 1255 TTTAAATTTTAGGAACCTTCAGAAATATCTTTTGAAGAGGTACTGCTAACTATAGTCAACCC 1314
 Db 1255 TTTAAATTTGGAGAAATCCCTGAATTTCTTT---AGAGGTAGCTTCTCTATCTATATAGG 1311
 QY 1315 TATGAGTCACTGGCTTCAATTAAGATTTAGAAATCTGAAATTTACCACAGAAACAAACA 1374
 Db 1312 TATACTGGAGTGGGACACAACTATTTGATTCAGAAATCTGAATTTACCACAGAAACAAACA 1371
 QY 1375 GAACGACCAAAATATGAATCATATAGTTCATAGGTATCTCACATAGGGCTCATTTCAAA 1434
 Db 1372 GAACGACCAAAATATGAATCTTACAGTCAATAGATTTATTAATTAAGACTAATATCAGGA 1431
 QY 1435 TCTAGGCTGATGTACCAGTATATTTCTTGACGACCGGTAGTCAGATCGTACAAATACC 1494
 Db 1432 AACACTTTGAGACCCAGTATATTTCTTGACGACCGGTAGTCAGATCGTACAAATACC 1491
 QY 1495 ATTAGTTTCAAGTACGATACACAAATACCATTTGGTAAATCATTTCAACCTTAAATTCAGGT 1554
 Db 1492 ATTAGTTTCAAGTACGATACACAAATACCATTTGGTAAATCATTTCAACCTTAAATTCAGGT 1551
 QY 1555 ACCTCTGTAGTGTGCTTCAAAACAAATGCTTGAAGGTAAATTTCCGAACTTAACTGTTAAT 1614
 Db 1552 ACCTCTGTAGTGTGCTTCAAAACAAATGCTTGAAGGTAAATTTCCGAACTTAACTGTTAAT 1611
 QY 1615 GGTAGTGTACTAAGTATGGGTCTTAAATTTTAAATTAATACATTTACAGCGGTATCGGCTG 1674
 Db 1612 GGTAGTGTACTAAGTATGGGTCTTAAATTTTAAATTAATACATTTACAGCGGTATCGGCTG 1671
 QY 1675 AGAGTTTGTATGCTGCTTCTCAAAACAAATGCTTGAAGGTAACTGTCGAGGGAGTACT 1734
 Db 1672 AGAGTTTGTATGCTGCTTCTCAAAACAAATGCTTGAAGGTAACTGTCGAGGGAGTACT 1731
 QY 1735 ACTTTTGATCAAGATTTCCCTAGTACTATGAGTGCATAATGAGTCTTTGACATCTCAATCA 1794
 Db 1732 ACTTTTGATCAAGATTTCCCTAGTACTATGAGTGCATAATGAGTCTTTGACATCTCAATCA 1791
 QY 1795 TTTAGATTTGCAAGATTTCCCTAGTACTATGAGTGCATAATGAGTGCATAATGAGTGCATA 1854
 Db 1792 TTTAGATTTGCAAGATTTCCCTAGTACTATGAGTGCATAATGAGTGCATAATGAGTGCATA 1851
 QY 1855 AGTATAGTAAATTAATGAGGTAGACAAACGTTTCACTTTGATAAAATTTGAATTCATTTCCA 1914
 Db 1852 AGTATAGTAAATTAATGAGGTAGACAAACGTTTCACTTTGATAAAATTTGAATTCATTTCCA 1911
 QY 1915 ATTAAGTCAACCTTCGAAGCAGAAATAGATTTTGAAGGGCGCAAGGGCGGTGAATGCT 1974
 Db 1912 ATTAAGTCAACCTTCGAAGCAGAAATAGATTTTGAAGGGCGCAAGGGCGGTGAATGCT 1971
 QY 1975 CTGTTTACTTAATACGATCCAGAGATTCGAAACAGATGTCAGATTTATCATTTGAT 2034
 Db 1972 CTGTTTACTTAATACGATCCAGAGGTTCGAAACAGATGTCAGATTTATCATTTGAT 2031
 QY 2035 CAAGTATCCAAATTTAGTGGGTGTTTATCGGATGAATTTCTGCTTAGATGAAAGAGAGAA 2094
 Db 2032 GAAGTATCCAAATTTAGTGGGTGTTTATCGGATGAATTTCTGCTTAGATGAAAGAGAGAA 2091
 QY 2095 TTAAGTGAAGAGTGAATATGCAAGATGCAAGATGCAAGATGCAAGATGCAAGATGCAAGAT 2154
 Db 2092 TTAAGTGAAGAGTGAATATGCAAGATGCAAGATGCAAGATGCAAGATGCAAGATGCAAGAT 2151
 QY 2155 CCAAACTTCACATCCATCAATAAGCAACAGACTTTCATATCTCTAATAGCAATCAAT 2214
 Db 2152 CCAAACTTCACATCCATCAATAAGCAACAGACTTTCATATCTAATAGCAATCAAT 2211


```

; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: entomocidus HD 110
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 186..3872
; OTHER INFORMATION: /note= "PROPERTIES: CryIB is toxic to
; OTHER INFORMATION: Ostrinea nubilalis (among others)"
US-08-377-690-1

Query Match      85.0%; Score 3133.6; DB 1; Length 4074;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 3373; Conservative 0; Mismatches 299; Indels 30; Gaps 3;

Qy 1 TTGACTTCAATAGGAAATAGAAATGAAATGAAATTAATAATGCTTTATCGATTCGAGCTGTA 60
Db 186 TTGACTTCAATAGGAAATAGAAATGAAATTAATAATGCTTTATCGATTCGAGCTGTA 230
Qy 61 TCGAATCATTCACACAAATGATCTATACCGATGCTCGTATTGAGGATTCCTTGTGT 120
Db 231 TCGAATCATTCGACAAATGATCTATACCGATGCTCGTATTGAGGATGAGCTTGTGT 290
Qy 121 ATAGCCGAGGGAAATATATCAATCCACTTGTGTAGCGCATCAACAGTCCAAACGGGTATT 180
Db 291 ATAGCCGAGGGAAATATATGATCCATTTGTGTAGCGCATCAACAGTCCAAACGGGTATT 350
Qy 181 AACATAGCTGTAGAAATCTAGTGTGATATAGCGGTAACGTTTGTGACAAATAGTAGT 240
Db 351 AACATAGCTGTAGAAATCTAGGCGTATTTGGCGTACCGTTTGTGACAAATAGTAGT 410
Qy 241 TTTTATAGTTTCTGTGTGTGATATGCGCGCGCGAGATCAGTGGGAAATTTTC 300
Db 411 TTTTATAGTTTCTGTGTGTGATATGCGCGCGCGAGATCAGTGGGAAATTTTC 470
Qy 301 CTAGAACATGTGCAACAACTTATAAATCAACAAATACAGAAATGCTAGGAAATCGGCA 360
Db 471 CTAGAACATGTGCAACAACTTATAAATCAACAAATACAGAAATGCTAGGAAATCGGCT 530
Qy 361 CTTGCTCGATTACAGGTTTAGGAGATTCCTTTAGAGCTATCAACAGTACCTTTGAAGAT 420
Db 531 CTTGCTCGATTACAGGTTTAGGAGATTCCTTTAGAGCTATCAACAGTACCTTTGAAGAT 590
Qy 421 TGGCTAGAAACCGTGTATGTCAGAACGAGAGTGTCTTTATACCAATATATAGCC 480
Db 591 TGGCTAGAAACCGTGTATGTCAGAACGAGAGTGTCTTTATACCAATATATAGCT 650
Qy 481 TTAGAACTTGATTTTCTTAATGCGATGCGGCTTTTCGCAATTAGAAACCAAGAGTTCCA 540
Db 651 TTAGAACTTGATTTTCTTAATGCGATGCGGCTTTTCGCAATTAGAAACCAAGAGTTCCA 710
Qy 541 TTATTAATGATATGCTCAAGCTGCAAAATTTACACCTATTTATTTAGAGATGCTCT 600
Db 711 TTATTAATGATATGCTCAAGCTGCAAAATTTACACCTATTTATTTAGAGATGCTCT 770
Qy 601 CTTTGTGATGATTTGGGCTTACATCGAGAAATTCGACGTTATTATGAGCGCAA 660
Db 771 CTTTGTGATGATTTGGGCTTACATCGAGAAATTCGACGTTATTATGAGCGCAA 830
Qy 661 GTGGAAACAAACGAGAGATTTATTCGACTATTTGCGTGTAGAAATGGTATAATACAGGTCTAAAT 720
Db 831 GTGGAAACGAGAGATTTATTCGACTATTTGCGTGTAGAAATGGTATAATACAGGTCTAAAT 890
Qy 721 AGCTTGAGGGCAAAATGCGCAAGTTGGGTGCGTTATATCAATTCGCTAGAGATCTA 780
Db 891 AGCTTGAGGGCAAAATGCGCAAGTTGGGTGCGTTATATCAATTCGCTAGAGATCTA 950
Qy 781 ACGTTAGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGACATCGCACCTTATCCA 840
Db 951 ACGTTAGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGACATCGCACCTTATCCA 1010
Qy 841 ATAAATACGAGTGTCTAGTTTAAAGGGAAGTTTATACAGACGCAATTTGGAGCAACAGGG 900

```

```

1011 ATAAATACGAGTGTCTCAGTTTAAACAAGAGAGTTTATACAGACGCAATTTGGAGCAACAGGG 1070
Qy
901 GTAAATATGCAAGTATGATTCGTAATAATAATGCACTTCCTCGTTTCCGCTATATAGAG 960
Db
1071 GTAAATATGCAAGTATGATTCGTAATAATAATGCACTTCCTCGTTTCCGCTATATAGAG 1130
Qy
961 ACTGCGGTTATCGGAAGCCCGCATCTACTTTGATTTTCTAGAAACAACCTTACAATTTTATG 1020
Db 1131 GCTGGGCTATCGGAAGCCCGCATCTACTTTGATTTTCTAGAAACAACCTTACAATTTTATG 1190
Qy
1021 ACTTCATCACGATGAGTGTCTACTAGGCAATGATGATTTACTGCGGGGGGCAACAATTCAA 1080
Db 1191 GCTTCATCACGATGAGTGTCTACTAGGCAATGATGATTTACTGCGGGGGGCAACAATTCAA 1250
Qy
1081 TCTCGGCAATAGGAGCGGATTAATAATACCTCAACGATCGGTCTACCAATCTCTTCTATT 1140
Db 1251 TCTCGGCAATAGGAGCGGATTAATAATACCTCAACGATCGGTCTACCAATCTCTTCTATT 1310
Qy
1141 AATCCTGTAAAGATTATCATTTCTCTCGAGAGCTATTTTGGAGTGAATCATATGCAAGGA 1200
Db 1311 AATCCTGTAAATTTACGTTTCGCATCTCGAGAGCTTTATAGGACTGAATCATATGCAAGGA 1370
Qy
1201 GTGCTTCTATGCGGAATTTACCTTGAACCTTATTCATGCTGCTCCCTACTGTTAGATTTAAT 1260
Db 1371 GTGCTTCTATGCGGAATTTACCTTGAACCTTATTCATGCTGCTCCCTACTGTTAGATTTAAT 1430
Qy
1261 TTTAGGAAACCTCAGATATCTTTTGAAGAGTACTGCTAACTATATAGTCAACCTATGAG 1320
Db 1431 TTTAGGAAACCTCAGATATCTTTTGAAGAGTACTGCTAACTATATAGTCAACCTATGAG 1490
Qy
1321 TCACCTGGGCTTCAATTTAAAGATTTCAAGAACTTGAATTTACCAACAGAAACCAACAAGACGA 1380
Db 1491 TCACCTGGGCTTCAATTTAAAGATTTCAAGAACTTGAATTTACCAACAGAAACCAACAAGACGA 1550
Qy
1381 CCAAAATATGAATCATATAGTATCATAGGTTATCTCACAATGAGGCTCATTTCAACATCTAGG 1440
Db 1551 CCAAAATATGAATCATATAGTATCATAGGTTATCTCACAATGAGGTTATTTTACAACTCCAGG 1610
Qy
1441 GTGATGATCCAGTATATCTTGGAGCGACCGTAGTGCAGATCGTACAAATACCATTAGT 1500
Db 1611 GTGATGATCCAGTATATCTTGGAGCGACCGTAGTGCAGATCGTACAAATACCATTAGT 1670
Qy
1501 TCAGATAGCATAAACAACAATACCAATTTGGTAAATTCATTTCAACCTTAATTCAGGTAACCTCT 1560
Db 1671 CCAAAATAGATCACCCAAATCCCAAATGTTAAAGCATCCGAACTTCCTCAAGGTACCACT 1730
Qy
1561 GTAGTACGTGCGCCAGGATTTACAGGAGGGATATAATCCGAACATAAGTCTTAATGAGT 1620
Db 1731 GTTGTAGAGGACGAGGATTTACTGTTGGGGGATATTTCTCGAAGAACGAATCTGTTGGGA 1790
Qy
1621 GTACTAGTATGGGTCTTAATTTTAAATAATACATCATTTACAGCGGTATCGCTGAGAGTT 1680
Db 1791 TTTGGACCGATAGAGTACTGTTTACGGACCATTAACAAGAGATATCGTATAGGATTC 1850
Qy
1681 CGTTATCTCTCTCTCAAAACAATGGTCTCTGAGGGTAACTGTGCGAGGGAGTACTACTTTT 1740
Db 1851 CGCTATGCTTCAACTGTAGATTTTGTATTTCTTGTATCACGTGGAGGTACTACTGTAAAT 1910
Qy
1741 GATCAAGGATTCCTCTAGTACTATAGTGTGCAAAATGAGTCTTTTGACATCTCAATCAATTTAGA 1800
Db 1911 AATTTTGAATTCCTTACGTAATGAACAGTGGAGACGAATTAATAATACGGAATTTTGTG 1970
Qy
1801 TTTGCAAAATTTCTGTAGGTATTTAGTGCATCTGGCAGTCAA --- ACTGCTGGAATAAGT 1857
Db 1971 AGAGTGTCTTTTACTACACCTTTTACTTTTACAAATTCAGAGATATAATTCGAACGTCT 2030
Qy
1858 ATAAGTAAATATGCAAGTATGCAAAACGTTTTCATTTGATATAAATTTGAATTCATTTCCAAT 1917
Db 2031 ATTCAGGCGCTTAGTGGAAATGGGAAGTGTATATAGATAAATTTGAATTTATTTCCAGTT 2090
Qy
1918 ACTGCAACCTTCGAGCAGAAATACGATTTTGAAGGGCGGCAAGAGCGGTGAATGCTCTG 1977
Db 2091 ACTGCAACCTTCGAGCAGAAATATGATTTTGAAGAGCGGCAAGAGCGGTGAATGCTCTG 2150

```

Qy	1978	TTTACTAATACGAATCCAAGAAGATTGAAACAGATGTGCACAGANTTACATATTGATCAA	2037
Db	2151	TTTACTAATACGAATCCAAGAAGATTGAAACAGATGTGCACAGATTATCATATTGATCAA	2210
Qy	2038	GTATCCAAATTAGTGGCGTGTTCATCCGATGCAATTCCTGCTTAGATGAAAGAGAGAATTA	2097
Db	2211	GTATCCAAATTAGTGGCGTGTTCATCCGATGCAATTCCTGCTTAGATGAAAGAGAGAATTA	2270
Qy	2098	CTTGAGAAAGTGAAATATGCGAAACGACTCAGTGCATGAAAGAAACTTTACTCCAGATCCA	2157
Db	2271	CTTGAGAAAGTGAAATATGCGAAACGACTCAGTGCATGAAAGAAACTTTACTCCAGATCCA	2330
Qy	2158	AACTTTCATCTCCATCAATTAAGCAACACGACTTCATATCTACTATAGCAATTCGAATTC	2217
Db	2331	AACTTTCATCTCCATCAATTAAGCAACACGACTTCATATCTACTATAGCAATTCGAATTC	2390
Qy	2218	ACATCTATCCATGAAACAATCTGAACTAGATGGTGGGGAAGTGAGAACATTTACAATCCAG	2277
Db	2391	ACATCTATCCATGAAACAATCTGAACTAGATGGTGGGGAAGTGAGAACATTTACAATCCAG	2450
Qy	2278	GAAGGAAATGACGTATTTAAAGAGAAATTACGTCACTACCGGGGACTTTTAATAGTGT	2337
Db	2451	GAAGGAAATGACGTATTTAAAGAGAAATTACGTCACTACCGGGTACTTTTAATAGTGT	2510
Qy	2338	TATCCGACGTATTTATATCAAAAAATAGGAGAGTCGGAATTTAAAGCTTATCTCGCTAC	2397
Db	2511	TATCCGACGTATTTATATCAAAAAATAGGAGAGTCGGAATTTAAAGCTTATCTCGCTAC	2570
Qy	2398	CAATTAAAGCGGTATATTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTATAAT	2457
Db	2571	CAATTAAAGCGGTATATTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTATAAT	2630
Qy	2458	GCAGAACATGAACATTTGATGTTTCCAGGTACCGAGTCCGATGCGCGCTTTTCAGTTGAA	2517
Db	2631	GCAGAACATGAACATTTGATGTTTCCAGGTACCGAGTCCCTATGCGCGCTTTTCAGTTGAA	2690
Qy	2518	AGCCCAATCCGAAGTCCGAGAGACCGAATCGATCGCACCAATTTTGAATCGGAATCCT	2577
Db	2691	AGCCCAATCCGAAGTCCGAGAGACCGAATCGATCGCACCAATTTTGAATCGGAATCCT	2750
Qy	2578	GATCTAGATTGTTCTCTCAGAGATGGAGAAAAATGTCCGCATCATTTCCCATCATTTCTCT	2637
Db	2751	GATCTAGATTGTTCTCTCAGAGATGGAGAAAAATGTCCGCATCATTTCCCATCATTTCTCT	2810
Qy	2638	TTGGATTTGANTTTGGATGCACGACTTCGATGAGAAATCTAGCGGTGGGTGGTATTC	2697
Db	2811	TTGGATTTGANTTTGGATGCACGACTTCGATGAGAAATCTAGCGGTGGGTGGTATTC	2870
Qy	2698	AAGATTAAAGACGCGAAGAGTCAATGCAAGACTAGGGAATCTGGAATTTATTGAAGAGAAA	2757
Db	2871	AAGATTAAAGACGCGAAGAGTCAATGCAAGACTAGGGAATCTGGAATTTATTGAAGAGAAA	2930
Qy	2758	CCATTATTAGGAGAACACTGTTCTGTTGGAAGAGCAGAGAGAAAAAATCGAGAGACAAA	2817
Db	2931	CCATTATTAGGAGAACACTGTTCTGTTGGAAGAGCAGAGAGAAAAAATCGAGAGACAAA	2990
Qy	2818	CGTGAAAACTACATTTGGAAAAAACAACGAGTATATACAGAGGCAAAAGACTGTGGAT	2877
Db	2991	CGTGAAAACTACATTTGGAAAAAACAACGAGTATATACAGAGGCAAAAGACTGTGGAT	3050
Qy	2878	GCTTTATTTCTGAGATCTTCAATATAATAGATTAACAGCGGATACAACATTCGCATGATT	2937
Db	3051	GCTTTATTTCTGAGATCTTCAATATGATAGATTTACAGCGGATACAACATTCGCATGATT	3110
Qy	2938	CATCGCGCAGATAAATCTGTTTCATCGAATTCGAGAGGCTTATCTGTACAGAAATTCCTGTT	2997
Db	3111	CATCGCGCAGATAAATCTGTTTCATCGAATTCGAGAGGCGTATCTTTCAGAAATTCCTGTT	3170
Qy	2998	ATCCGGGTGTAAATGCCGAAAAATTTTGAAGAAATTAGAAGGTGCATATTCACCTGCATC	3057
Db	3171	ATCCCGAGGTGTAAATGCCGAAAAATTTTGAAGAAATTAGAAGGTGCATATTCACCTGCATC	3230

Qy	3058	TCCCTATACGATCGGAGAAATGTCGTTAAAAATGGTGATTTTAAATAAATGGATTAGCATGC	3117
Db	3231	TCCCTTATACGATCGGAGAAATGTCGTTAAAAATGGTGATTTTAAATAAATGGATTAAACATGT	3290
Qy	3118	TGGAATGTAAAGGGCATGTAGATGTACAAACAGAGCCATCACCGTTCGTGCTTGTGTTATC	3177
Db	3291	TGGAATGTAAAGGGCATGTAGATGTACAAACAGAGCCATCATCGTTCTGACCTTGTGTTATC	3350
Qy	3178	CCAGAAATGGGAAGCAGAAAGTGTCCAAAGCAGTTCGCGTCTGTCCGGGGCGTGCGCTATATC	3237
Db	3351	CCAGAAATGGGAAGCAGAAAGTGTCCAAAGCAGTTCGCGTCTGTCCGGGGTGTGGCTATATC	3410
Qy	3238	CTCCGCTGCACAGGTCACAAAGGCGATATGGAGAGGGTTGTGTACGATCCATGAAATC	3297
Db	3411	CTTCGCTGCACAGGTCACAAAGGCGATATGGAGAGGGCTGCGTAACGATCCATGAAATC	3470
Qy	3298	GAGAAACAATACAGACGAACTTAAAAATTTAAAAACTGTGAAAGAGAGGAAAGTGTATCCAAAG	3357
Db	3471	GAGAAACAATACAGACGAACTTAAAAATTTAAAAACCGTGAAGAGAGGAAAGTGTATCCAAAG	3530
Qy	3358	GATACAGGAACGTTGAATGATTATATCTGCACACCAAGGTACAGC-----AGTA	3405
Db	3531	GATACAGGAACGTTGAATGATTATATCTGCACACCAAGGTACAGCTGGATCGCGAGATGCA	3590
Qy	3406	TGTAATTCGCGTAACTGCTGGATATAGGATGCAATATGAAGTTGATACTACAGCATCTGTT	3465
Db	3591	TGTAATTCGCGTAACTGCTGGATATAGGATGCAATATGAAGTTGATACTACAGCATCTGTT	3650
Qy	3466	AATTACAAACCGACTTTATGAAGAAGAAACGTTATACAGATGTACGAAGAGATAATCATTTGT	3525
Db	3651	AATTACAAACCGACTTTATGAAGAAGAAACGTTATACAGATGTGAAGAAGATATCATTTGT	3710
Qy	3526	GAATATGACAGGGGTATGTGAATTATCCACNCTACAGCTGGTTATATGACAAAGAA	3585
Db	3711	GAATATGACAGGGGTATGTCAATTATCCACCAAGTACCAGCTGGTTATGTGCAAAAGAA	3770
Qy	3586	TTAGAATACTTCCAGAAACCGATAAGTGTATGGATTGAGATTGGAGAAACCGAAGGGAAG	3645
Db	3771	TTAGAATACTTCCAGAAACCAAGTATACAGTATGGATTGAGATTGGAGAAACCGNAGGAAG	3830
Qy	3646	TTTATTTGTAGACAGCGTGAATTTACTTCCTTTATGGAGGAATAG	3687
Db	3831	TTTATTTGTAGTAGCGTGAATTTACTTCCTCATGGAGGAATAG	3872

CLASSIFICATION: 514		Query Match		DB 1; Length 3934;	
ATTORNEY/AGENT INFORMATION:		Best Local Similarity 83.5%; Pred. No. 0;		Matches 3088; Conservative 0; Mismatches 593; Indels 15; Gaps 3	
NAME: Egolf, Christopher	REGISTRATION NUMBER: 27633	REFERENCE/DOCKET NUMBER: 7205-49	TELEPHONE: 215-757-1590	INFORMATION FOR SEQ ID NO: 3:	SEQUENCE CHARACTERISTICS:
LENGTH: 3934 base pairs	TYPE: nucleic acid	STRANDEDNESS: double	TOPOLOGY: circular	MOLECULE TYPE: DNA (genomic)	FEATURE:
NAME/KEY: CDS	LOCATION: 67..3756	NAME/KEY: misc feature	LOCATION: 2253..2272	US-08-100-709-3	
QY	1	TTGACTTCAATAGGAAAAATGAGAAATCAAAATTAATAATGCTTTATCAATCCAGCTCTGA	60		
DB	67	TTGACTTCAATAGGAAAAATGAGAAATCAAAATTAATAATGCTTTATCAATCCAGCTCTGA	126		
QY	61	TCGAATCAATCCACACAAATGGATCTATCACAGATGCTCGTATTGAGGATTCCTTTGTGT	120		
DB	127	TCGAATCCTTCCACGAAATGAATCTATCACAGATGCTCGTATTGAGGATTCCTTTGTGT	186		
QY	121	ATAGCCGAGGGAATATATCAATCCACTTGTATTAGCGCATCAACAGTCCAAACGGGTATT	180		
DB	187	GTAGCCGAGGGAATATATCAATCCACTTGTATTAGCGCATCAACAGTCCAAACGGGTATA	246		
QY	181	AACATAGCTGTAGAACTACTAGGTGTATTAGCGGTACCGTTTGTCTGGACAAATAGCTAGT	240		
DB	247	AACATAGCTGTAGAACTACTAGGTGTATGGCGGTATTAGGTGTGCGTTGCTGGCACTAGT	306		
QY	241	TTTTATAGTTTTCTTTGTGGTGAATATTAGCCCGCGCGCAGAGATCAGTGGGAAATTTTC	300		
DB	307	TTTTATAGTTTTCTTTGTGGGAAATATTAGCGCTAGTGGCAGAGATCCATGGGAAATTTTC	366		
QY	301	CTAGACATGTCGAACAACTTATTAATCAACAAATACAGAAATGCTAGGAATACGGCA	360		
DB	367	CTGGAAATGTAGAACAACTTATTAAGAACAAAGTAACAGAAATACTAGGAATACGGCT	426		
QY	361	CTTGCTCGATTACAAGGTTTAGGAGATTCTTTTAGCGCTATCAACAGTCACTTTGAAGAT	420		
DB	427	ATTGCTCGATTAGAAGGCTTAGGAAGAGGCTATAGATCTTACCAGCAGGCTCTTGAACT	486		
QY	421	TGGCTAGAAACCGGTGATGATGCAAGAACGAGAGAGTGTCTTTATACCAATATATAGCC	480		
DB	487	TGCTTAGATAACCGAAATGATGCAAGATCAAGAAGCATTTCTTTGAGCGCTATGTTGCT	546		
QY	481	TTAGAACTTGATTTTCTTAATGCGATGCGCTTTTTCGCAATTTAGAAACCAAGAAGTTCCA	540		
DB	547	TTAGAACTTGACATTAATACTGCTATACCGCTTTTCAGAAATACGAAATGAAGAATTTCCA	606		
QY	541	TTATTAAATGATATGCTCAAGCTGCAAAATTTACCTATTATTATTGAGAGATGCCTCT	600		
DB	607	TTATTAAATGATATGCTCAAGCTGCAAAATTTACCTATTATTATTGAGAGATGCCTCT	666		
QY	601	CTTTTTCGTAGTGAATTTGGGCTTTACATCGCAGGAAATTCACGTTATTATGAGCGCAA	660		
DB	667	CTTTTTCGTAGTGAATTTGGGGAATGGCATCTTCCGATGTTAACCAATATTACAGAACAA	726		
QY	661	GTGGAAACAAACGAGAGATTATTCGACCTATTTCGCTAGAAATGGTATATAACAGGCTAAAT	720		
DB	727	ATCAGATATACAGAGGAATATTTCAACCATTCGCTACAAATGGTATATAACAGGCTAAAT	786		

QY 1792 TCATTAGATTGGAGAAATTTCTCTAGGTATTAGTGCATCTGGCAGTCAAACTGCTGGA 1851
Db 1861 AGTTTTAGAACTGCGAGGATTTAGTACTCTCTTTTAAATTTGCCCCAAAGCACATTC 1920
QY 1852 ATAAGTATAAGTAATAATGCGAGGTAGACAAAGTTTCACTTTGATAAAATTTGAATTCATT 1911
Db 1921 ACATTTGGGTGCTCAGAGTTTTCATCAATCAGGAAGTTTATATAGATAGAGTCAATTTGTT 1980
QY 1912 CCAATTTACTGCAACTTCGAAAGCAGAAATACGATTTAGAAAGGCGCGCAAGAGCGGTGAAT 1971
Db 1981 CCAGCAGAGGTAAACATTTTGAGCGCAGAAATATGATTTAGAAAGAGCACAAAGCGCGTGAAT 2040
QY 1972 GCTCTGTTTACTAATACCAATCCAAAGAGATGAAACAGATGTGACAGATATCATATT 2031
Db 2041 GCTCTGTTTACTTCTACAAATCCAAAGAGATGAAACAGATGTGACAGATATCATATT 2100
QY 2032 GATCAAGTATCAATTTAGTGGCGTGTATTCGGATGAATTTCTGTTAGATGAAAGAGA 2091
Db 2101 GACCAAGTGTCAATATGCTGGCATGTTTATCAGATGAATTTTCTGATGAGAGCGA 2160
QY 2092 GAATTTACTTGAGAAAGTGAATATGCGAAACGACTCAGTGATGAAAGAACTTTACTCCAA 2151
Db 2161 GAATTTATTTGAGAAAGTGAATATGCGAAAGCAGTCACTGATGAAAGAACTTTACTCCAA 2220
QY 2152 GATCCAACTTTCATCCATCAATTAAGCAACAGAGCTTTCATCTACTAATGAGCAATCG 2211
Db 2221 GATCCAACTTTCATCCATCACTGAGTGGCAATTAAGTTTTCGATCCATCGATGGAATCA 2280
QY 2212 AATTTACATCTATCTCAATGAACAATCTGAACATGATGCTGGGAAGTGAGAAACATTACA 2271
Db 2281 AACTTCCCTCTATTAATGAGCTATCTGAAATGATGATGCTGGGAAGTGCGAATGTTTACC 2340
QY 2272 ATCCAGGAAGGAATGAGCTATTTAAAGAGAAATTAAGTCACTACCGGGGACTTTTAAAT 2331
Db 2341 ATTCCAGGAAGGAATGAGCTATTTAAAGAGAAATTAAGTCACTACCGGGTACTTTTAAAT 2400
QY 2332 GAGTGTATCCAGCTATTTATATCAAAATATAGAGAGTGGAAATTAAGAGCTTTACT 2391
Db 2401 GAGTGTATTCAAATTTATTTATCAAAATATAGAGAGTGGAAATTAAGAGCTTTACTACG 2460
QY 2392 CGCTACCAATTAAGAGGCTATTTGAGATAGTCAAGATTTAGAGATATTTTCACTTCGT 2451
Db 2461 CGCTATCAATTAAGAGGCTATTTGAGATAGTCAAGATTTAGAGATTTTAAATTCGT 2520
QY 2452 TATTAATGCGAAACATGAAACATTTGGATGTTCCAGGTACCGAGTCCGTATGCGCGCTTTCA 2511
Db 2521 TACATGCAAGCATGAAACATTTGGATGTTCCAGGTACCGATTTCCCTATGCGCGCTTTCA 2580
QY 2512 GTTGAAAGCCCAATCGGAAGGTGCGGAGAACCGAATCGATCGGCAACCAATTTTGAATGG 2571
Db 2581 GTTGAAAGCCCAATCGGAAGGTGCGGAGAACCAATCGATCGGCAACCAATTTTGAATGG 2640
QY 2572 AATCTGATCTAGATGTTCTGTCAGAGATGGAAGAAATGTCGCGCATCTTCCCATCAT 2631
Db 2641 AATCTGATCTAGATGTTCTGTCAGAGATGGAAGAAATGTCGCGCATCTTCCCATCAT 2700
QY 2632 TTCTCTTTGGAATTTGATTTGATGTCACAGACTTTGCATGAGAATCTAGGCGTGTGGGTG 2691
Db 2701 TTCACTTTGGATTTGATTTGGTGGCAGACTTTGCATGAGACCTTAGGCGTGTGGGTG 2760
QY 2692 GTATTCAGATTAAGACGAGAGGTGTCATCAAGATAGTAGGAAATCTGGAATTTTATGAA 2751
Db 2761 GTATTCAGATTAAGACGAGAGGTGTCATCAAGATTTAGGAAATCTGGAATTTTATCGAA 2820
QY 2752 GAGAAACATTTATAGGAGAGCTGCTGCTGTGAGAGAGCAGAGAGAGAGAGAGAGAGAGAG 2811
Db 2821 GAGAAACATTTATTTGAGAGAGCTGCTGCTGTGAGAGAGCAGAGAGAGAGAGAGAGAGAG 2880
QY 2812 GACAAACGTGAAAACTCAATTTGAAACAAAAACGATATATACAGAGGCGCAAAAGAGCT 2871
Db 2881 GACAAACGTGAAAACTCAATTTGAAACAAAAACGATATATACAGAGGCGCAAAAGAGCT 2940
QY 2872 GTGGATGCTTTTATTTGTAGATTCTCAATATATAGATTACAAAGCGGATACAAACATTTGGC 2931

Db 2941 GTGGATGCTTTTATTTGTAGATTCTCAATATGATCAATTTACAAGCGGATACAAACATTTGGC 3000
QY 2932 ATGATTCATGCGGAGAGATAAACTTTGTTTCATCGAATTTGAGAGGCTTATCTGTCAAGATTA 2991
Db 3001 ATGATTCATGCGGAGAGATAAACTTTGTTTCATCGAATTTGAGAGGCTTATCTTTCAGATTA 3060
QY 2992 TCTGTTATCCCGGTGTAATTCGCGAAATTTTGAAGAAATTTAGAAAGTTCGATTTACTACT 3051
Db 3061 CTTGTTATCCCGGTGTAATTTGCGGAAATTTTGAAGAAATTTAGAAAGTTCATTTACTACT 3120
QY 3052 GCATCTCCCTATACGATGCGAGAAATGTCGTTAAAAATGCTGATTTTAAATTTGATTA 3111
Db 3121 GCATGCTCTTATACGATGCGAGAAATGTCGTTAAAAATGCTGATTTTAAATTTGATTA 3180
QY 3112 GCATGCTGGAATGTTAAAAAGGCGATGTAGATGTACAAAGAGGCTTGTGTAACGATTCAT 3171
Db 3181 ACATGTTGGAATGTTAAAAAGGCGATGTAGATGTACAAAGAGGCTTGTGTAACGATTCAT 3240
QY 3172 GTTATCCAGAAATCGGAAGCAGAAAGTGTCAAGAGTTCGCGTCTGTCGCGGGCGTGGC 3231
Db 3241 GTTATCCAGAAATCGGAAGCAGAAAGTGTCAAGAGTTCGCGTCTGTCGCGGGCGTGGC 3300
QY 3232 TATATCTCTCGTGTACAGCGTACAAAGAGGATGTGAGAGGCTTGTGTAACGATTCAT 3291
Db 3301 TATATCTCTCGTGTACAGCGTACAAAGAGGATGTGAGAGGCTTGTGTAACGATTCAT 3360
QY 3292 GAAATCCAGAAACATACAGAGCAATTAATTTAAAACTGTGAAGAGAGAGAGTGTAT 3351
Db 3361 GAAATCCAGAAACATACAGAGCAATTAATTTAAAACTGTGAAGAGAGAGAGTGTAT 3420
QY 3352 CCAACCGATACAGAAACGTTGATTTATATCTGCACACCAAGGTACAGCATTTGTAAT 3411
Db 3421 CCAACCGATACAGAAACGTTGATTTATATCTGCACACCAAGGTACAGCATTTGTAAT 3480
QY 3412 TCCCGTAACTGCTGATGAGGATGCATATGAAGTTGATGATGATGATGATGATGATGAT 3471
Db 3481 TCCCGTAACTGCTGATGAGGATGCATATGAAGTTGATGATGATGATGATGATGATGAT 3540
QY 3472 AAACCGACTTATGAAGAGAAACGTTATACAGATGTACAGAGATTAATCATTTGTAAT 3531
Db 3541 AAACCGACTTATGAAGAGAAACGTTATACAGATGTACAGAGATTAATCATTTGTAAT 3600
QY 3532 GACAGAGGTATGTAATTTATCCACCTACCTGCTGTTATGATGATGATGATGATGATGATGAT 3591
Db 3601 GACAGAGGTATGTAATTTATCCACCTACCTGCTGTTATGATGATGATGATGATGATGATGAT 3660
QY 3592 TACTTCCAGAAACCGATGAGTATGAGATTTGAGATTTGAGAGAGAGAGAGAGAGAGAGAG 3651
Db 3661 TACTTCCAGAAACCGATGAGTATGAGATTTGAGATTTGAGAGAGAGAGAGAGAGAGAGAG 3720
QY 3652 GTAGACAGCGTGAATTTACTCTTATGAGAGAAATAG 3687
Db 3721 GTAGTACGCTGGAACCTACTCTCATGGAAGATAG 3756

RESULT 7

US-08-176-865-3
; Sequence 3, Application US/08176865
; Patent No. 5616319
; GENERAL INFORMATION:
; APPLICANT: Donovan, William P.
; APPLICANT: Tan, Yiping
; APPLICANT: Jany, Christine S.
; APPLICANT: Gonzalez Jr., Jose M.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
; ADDRESSEE: Nadel
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia

STATE: Pennsylvania
 COUNTRY: U.S.A.
 ZIP: 19103
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk.
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/176.865
 FILING DATE: 30-DEC-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/100,709
 FILING DATE: 29-JUL-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Egolf, Christopher
 REGISTRATION NUMBER: 27633
 REFERENCE/DOCKET NUMBER: 7205-49
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 215-757-1590
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3934 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 67..3756
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 2253..2272
 US-08-176-865-3

Query Match 72.9%; Score 2687.2; DB 1; Length 3934;
 Best Local Similarity 83.5%; Pred. No. 0;
 Matches 3088; Conservative 0; Mismatches 593; Indels 15; Gaps 3;

Qy	1	TTGACTTCAATAGAGAAATGAGAATCAAAATTAATAATGCTTTTATCGATTCCAGCTGTA	60
Db	67	TTGACTTCAATAGAGAAATGAGAATCAAAATTAATAATGCTTTTATCGATTCCAAACGGTA	126
Qy	61	TCGAATCAATCCACAAATGGATCTATCACCAGATGCTGCTATGAGGATCTTTGCT	120
Db	127	TCGAATCCTTCCAGCAAAATGAATCTATCACCAGATGCTGCTATGAGGATCTTTGCT	186
Qy	121	ATAGCCGAGGGGAATATATCAATCCACTTGTAGCGCATCAACAGTCCAAACGGTATT	180
Db	187	GTAGCCGAGGGGAATATATCAATCCACTTGTAGCGCATCAACAGTCCAAACGGGTATA	246
Qy	181	AACATAGCTGTAGAACTACTAGGTGTATAGGCGTA	240
Db	247	AACATAGCTGTAGAACTACTAGGTGTATAGGCGTATAGGTGCGGTGCTGGCAACTAGCTAGT	306
Qy	241	TTTTATAGTTTTCTTTGTTGGTAATATATGGCCCGCGCGCAGAGATCAGTGGGAATTTTC	300
Db	307	TTTTATAGTTTTCTTTGTTGGGAATATATGGCCTAGTGGCAGAGATCCATGGGAAATTTTC	366
Qy	301	CTAGAACATGTCGAACAACTTATTAATCAACAAATAACAGAAAATGCTAGGAATACGGCA	360
Db	367	CTGGAAACATGTAACAACTTATTAAGCAACAGTAAACAGAAATATCTAGGAAATACGGCT	426
Qy	361	CTTGCTCGAATTAAGGTTTATAGGAGATTCCCTTTAGAGCCTATCAACAGTCACTTGAAGAT	420
Db	427	ATTGCTCGAATTAAGGTTTATAGGAGAGGCTATAGATCTTACCAGCAGGCTCTTGGAACT	486
Qy	421	TGGCTAGAAACCGGTGATGATGCAAGAACGAGAGTGTCTTTTATACCAATATATAGCC	480
Db	487	TGGTTAGATAACCGAAATGATGCAAGATCAAGAGCAATTAATCTTGAGCGCTATGTTGCT	546
Qy	481	TTAGAACTTGATTTCTTTAATGCGATGCGGCTTTTGGCAATTAGAAACCAAGAGTTCCA	540

Db	547	TTAGAACTTGACATTACTACTGCTATACCGCTTTTCAGAAATACGAAATGAAGAAGTTCCA	606
Qy	541	TTATTAAATGCTATATGCTCAAGCTGCAAAATTTACACCTATATATTATGAGAGTGCCTCT	600
Db	607	TTATTAAATGCTATATGCTCAAGCTGCAAAATTTACACCTATATATTATGAGAGTGCCTCT	666
Qy	601	CTTTTGTAGTAGTGAATTTGGGCTTTACATCGCAGGAAATTCACGTTATTATGAGCGCAA	660
Db	667	CTTTTGTAGTAGTGAATTTGGGGAATGCGCATCTCCGATGTTAACCAATATTACCAAGAACAA	726
Qy	661	GTGGAACAAACGAGAGATTATTCGACTATTCGCTAGAAATGGTATAATACAGGTTCTAAAT	720
Db	727	ATCAGATATACAGAGGAATATTTCAACCAATGCGTACAAATGGTATAATACAGGCTTAAAT	786
Qy	721	AGCTTGAGAGGACAAATGCGCAAGTTGGTGGCTTATTAATCAATTCCTGAGAGATCTA	780
Db	787	AACCTTAAGAGGGACAAATGCTGAAAGTTGGTGGGTATAATCAATTCCTGAGAGATCTA	846
Qy	781	ACGTTAGGGGTATTAGATCTAGTGGCACTATTCCTCAAGCTATGACACTCGCAGCTTATCCA	840
Db	847	ACGTTAGGGGTATTAGATCTAGTGGCACTATTCCTCAAGCTATGATCTCGCACTTATCCA	906
Qy	841	ATAAATACGAGTGTCTCAGTTTAAACAAGGGAAGTTTATACAGCGCAATTTGGAGCAACAGGG	900
Db	907	ATCAATACGAGTGTCTCAGTTTAAACAAGGGAAGTTTATACAGATCCAAATGGGAGAACAAAT	966
Qy	901	GTAAT-----ATGCAAGTATGAATTTGGTATAATTAATGACACCTTGTTCGCT	954
Db	967	GCACTTTCAAGTATGCAAGTATGCAATTTGGTATAATTAATGACACCTTGTTCGCT	1026
Qy	955	ATAGAGACTCGGTTATCCGAAGCCCGCATCTACTTCAATTTCTTCTAGAACAACTTACAAAT	1014
Db	1027	ATAGAGCTGCCATTTTCAGGCTCCGCATCTACTTCAATTTCTTCTAGAACAACTTACAAAT	1086
Qy	1015	TTTAGCACTTTCATCAGCATGAGTGTACTAGGCATATGACTTACTCGCGGGGCAACACA	1074
Db	1087	TACAGTGCATCAAGCGCTTGGAGTAGCACTCAACATATGAATATTATGGTGGGACATAGG	1146
Qy	1075	ATTCAATCTCGGCAATAGAGGGCGGATTAATACCTCAACGGCATGGTCTACCA--	1131
Db	1147	CTTAACCTTCGCGCAATAGAGGGGACATTAATACCTCAACAGGACCTTACTTAATAT	1206
Qy	1132	ACTTCTATTAATCTCTGAAGATTATCATCTCTCTCGAGACGTATATTGGACCTGAATCA	1191
Db	1207	ACTTCAATTAATCTCTGTAACATTAACAGTTTACGCTCTCGAGACGTATTATAGAACAGATCA	1266
Qy	1192	TATGACAGAGTGTCTTATGGGGAATTTA	1251
Db	1267	AATGACAGGACAAATAT-----ACTATTTACTACTCTGTGAATGGAGTACCTTGGGCT	1320
Qy	1252	AGATTTAATTTTAGGAACCTCTCAGAACTATTTTGAAGAGGATCTGCTTAATATAGTCAA	1311
Db	1321	AGATTTAATTTTAAACCTCTCAGAAATTTATGAAGAGGCGCCATACCTACAGTCAA	1380
Qy	1312	CCCTATCAGTCACCTCGGCTTCAATTAAGAACTGAGAACTGAATTAACCAAGAACAA	1371
Db	1381	CCGTATCAGGAGTTGGGATTCATTTATTTGATTCAGAACTGAATTTACCAACGAGAACAA	1440
Qy	1372	ACAGAAGCAACCAATTAATGAATCATATAGTCATAGGTTATCTCATATAGGCTCATTTCA	1431
Db	1441	ACAGAAGCAACCAATTAATGAATCATATAGTCATAGTATCTCATATAGGACTAATCATAT	1500
Qy	1432	CAATCTAGGCTGATGATACCAAGTATATTTCTTGGAGCAGCCGCTAGTGCAGATCGTACAAAT	1491
Db	1501	GGAAACACTTTGAGAGCAGCAGTCTATTTCTTGGAGCAGTCTGATGAGATCGTACGAAAT	1560
Qy	1492	ACCATTAGTTTTCAGATAGCAACAAATACCATTTGGTAAATCATTTCAACCTTAAATTC	1551
Db	1561	ACGATTGCAACCAATAGAAATTAACAATACCATTTGGTAAAGACACTGAATCTTCATTC	1620
Qy	1552	GGTACCTCTGTAGTCACTGGCCCAAGGATTTACAGAGGGGATTAATTCGGAACCTAACGTT	1611

Db	1621	GGTGTTACTGTTGTTGGAGGCCAGGATTTACAGGTGGGGANATCCCTTGGTAGAACAAAT	1681
Qy	1612	AATGGTAGTGACTAAGATATGGGTCTTAAATTTTAAATAATCATCATTTACAGCGGTATCGC	1671
Db	1681	ACGGGTACATTTGGAGATATACGATTAATATTAATGTGCCATTTATCCCAAAGATATCGC	1740
Qy	1672	GTGAGAGTTTCGTTATGCTGCTTCTCAAACAAATGCTCTGAGGGTAACTGTCCGAGGGAGT	1731
Db	1741	GTAAGGATTCGTTATGCTTCTACTACAGATTTACAAATTTTTCACGAGAAATTAATGGGAACC	1800
Qy	1732	ACTACTTTTGTATCAAGGATTCCTCTAGTACTATAGTGCANAATGAGTCTTTCGACATCTCAA	1791
Db	1801	ACTGTTAAATTTGGTAATTTCTCAAGAACTATGAATAGGGGGGATAATTTAGAAATATAGA	1860
Qy	1792	TCATTTAGATTTGCAGAAATTTCTGTAGTATTAGTGCATCTGCGCAGTCAAACTGCTGGA	1851
Db	1861	AGTTTTAGAACTGCAGGATTTTAGTACTCTTTTAAATTTTTTAAATGCCAAAGCACATTC	1920
Qy	1852	ATAAGTATATAGTAATAATGCAAGGTAGACAAACGTTTCCTTTGTATAAAAATGTAATTCAT	1911
Db	1921	ACATTTGGGTGCTCAGAGTTTTCAATCAGGAAGTTTATATAGATAGAGTCGAATTTGTT	1980
Qy	1912	CCAATTACTGCAACCTTCGAGCAGAAATACGATTTAGAAAGGGCGCAAGAGGGCGGTGAAT	1971
Db	1981	CCAGCAGAGGTAACTTTGAGGCGAGATATGATTTAGANAAGCACAAAAGGCGGTGAAT	2040
Qy	1972	GCTCTGTTTACTAATACGAATCCAAAGAAGATTGAAAAACAGATGTGACAGATTTATCATATT	2031
Db	2041	GCCTCTGTTTACTCTACAAATCCAAAGAAGATTGAAAAACAGATGTGACAGATTTATCATATT	2100
Qy	2032	GATCAAGTATCCAAATTTAGTGGCGTGTATTACGGATGAAATCTCTCTAGATGAAAAGAGA	2091
Db	2101	GACCAAGTGCCTAATATGGTGGCATGTTTATCAGATGAAATTTTGTCTGGATGAGAAGCGA	2160
Qy	2092	GAATTACTTGAGAAAGTGAATAATGCGAAACGACTCAGTCAGTGAAGAAACAACTTACTCCAA	2151
Db	2161	GAATTTATTTGAGAAAGTGAATAATGCGAAGCGACTCAGTCAGTGAAGAAACAACTTACTCCAA	2220
Qy	2152	GATCCAAACTTCACATCCATCAATTAAGCAACACAGACTTCATATCTACTAATGAGCAATCG	2211
Db	2221	GATCCAAACTTCACATTCATCAGTGGCGCAATTAAGTTTCGCATCCATGAGCAATCA	2280
Qy	2212	AATTTACATCTATPCATGAACAATCTGAACATGGATGGTGGGGAAGTGAGAACATTTACA	2271
Db	2281	AACCTCCCTCTATTAAATGAGTATCTGAACATGGATGGTGGGGAAGTGGCAATTTTACC	2340
Qy	2272	ATCAGGAAGGAATGACGTATTTAAAGAGAAATTAAGTCACTACACGCGGGACCTTTTAAT	2331
Db	2341	ATTGAGGAAGGAATGACGTATTTAAAGAGAAATTAAGTCACTACACGCGGTACTTTTAAT	2400
Qy	2332	GAGTGTATTCGACGTATTTATCAAAAAATAGGAGAGTCGGAATTTAAAGCTTATACT	2391
Db	2401	GAGTGTATTCGAAATTTATTTATATCAAAAAATAGGAGAGTCAGAAATTAAGGCTTATACG	2460
Qy	2392	CGCTACCAATTTAAGAGGGTATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGT	2451
Db	2461	CGCTATCAATTAAGAGGGTATTTGAAGATAGTCAAGATCTAGAGATTTATTTAAATTCGT	2520
Qy	2452	TATTAATCGGAAAATANGAAAATTTGGATGTTCAGAGTTACCGAGTCCGATAGGCGCTTTCA	2511
Db	2521	TACAAATGCAAGCATGAAACATTTGGATGTTTCCAGGTACCGAATTCCTTATGGCGCTTTCA	2580
Qy	2512	GTTGAAGCCCAATCGAAGGTGCGGAGAACCGAATCGATGCCNCCACAAATTTGATATGG	2571
Db	2581	GTTGAAGCCCAATTCGAAGGTGCGGAGAACCAAAATCGATGCGCACCAATTTTGATATGG	2640
Qy	2572	AATCCTGATCTAGATTTGTTCTGCGAGAGATGGAGAAAAAATGTGCGCATCATTTCCCATCAT	2631
Db	2641	AATCCTGATCTAGATTTGTTCTGCGAGAGATGGAGAAAGATGTGGCATCATTTCCCATCAT	2700
Qy	2632	TTCTCTTTTGATATTTGATATTTGATGCAACAGACTTGCATCAGAAATCTAGCGGTGTGGTG	2691
Db	2701	TTCACTTTGATATTTGATTTGGGTGCGACAGACTTGCATGAGAACTCTAGCGGTGTGGTG	2760

RESULT 8
US-08-474-038-3
; Sequence 3, Application US/08474038

DEPT 6

RESULTS 8
IIS-08-474-038-3

: Sequence 3, Application US/08474038

—

Db 1381 CGTATCAGGAGTTGGGATTCATTAATTGATTTCAGAAACTGAATTAACCAAGCAACA 1440
Qy 1372 ACAGAACGACAAATTAATGAATCATATAGTCATAGGTTATCTCACATAGGGCTCATTTCA 1431
Db 1441 ACAGAACGACAAATTAATGAATCATATAGTCATAGGTTATCTCATATAGGACTAATCAT 1500
Qy 1432 CAATCTAGGGTCATGTACCAAGTATATTTCTGGAGCGACCGTAGTGAGATCGTACAAT 1491
Db 1501 GGAACACATTTTGAGAGCACCAAGTCTATTTCTGGACGCATCGTAGTGCAATCGTACGA 1560
Qy 1492 ACCATTAGTTCAGATGAGATACACAAATACCATTTGTTAAATCATTTCAACTTAATTTCA 1551
Db 1561 ACGATTGACCAATAGAAATACACAAATACCATTTGTTAAAGCACTGAATCTTCAITTC 1620
Qy 1552 GGTACCTCTGATGACGTGGCCAGGATTTACAGAGGGGATATAATCCGAACCTAACGTT 1611
Db 1621 GGTGTTACTGTTGAGAGGCGCAGGATTTACAGGTGGGATATCCTTCGTAGAACAAT 1680
Qy 1612 AATGTTAGTGTACTAAGTATGGGTCCTTAATTTTAATAATCATNTTACAGGGTATCGC 1671
Db 1681 ACGGTACATTTGGAGATATACGATTAATTAATGTCCTATTTATCCCAAGATATCGC 1740
Qy 1672 GTGAGAGTTCGTTATGCTGCTTCTCAACAAATGCTGAGGGTAACTGTGCGAGGGAGT 1731
Db 1741 GTAAGGATTCGTTATGCTTCTACTACAGATTTACAAATTTTCAAGAAATTAATGGAACC 1800
Qy 1732 ACTACTTTTGTACAAAGATTTCCCTAGTACTAGTGCAAAATGAGTCTTTGACATCTCAA 1791
Db 1801 ACTGTTAATATTTGTTAAATTTCTCAAGAACTATGAATAGGGGGATTAATTTAGATATAGA 1860
Qy 1792 TCATTTAGATTTGCAGAAATTTCTGTTAGGTATAGTGCAATCTGGCAGTCAAACTGTGGA 1851
Db 1861 AGTTTTAGAACTGCGAGGATTTAGTACTCTCTTTTAATTTTAAATGCCCAAGCACATTC 1920
Qy 1852 ATAAGTATAAGTATAATGTCAGGTAGACAAACGTTTCTACTTTGATAAAATGAATTCATT 1911
Db 1921 ACATGGGTGCTCAGAGTTTTTCAATCAGGAATTTATATAGATAGAGTCGAATTTGTT 1980
Qy 1912 CCAATTTACTGCAACCTTCGAAGCAGAAATACGATTTAGAAAGGGCGCAAGCGGTGAAT 1971
Db 1981 CCAGCAGAGGTAAACATTTGAGGCAGAAATATGATTTAGAAAGGCACAAAGCGGTGAAT 2040
Qy 1972 GCTCTGTTTACTAATPACGAATCCAAAGAGATGAAACAGATGTGACAGATTTATCATATT 2031
Db 2041 GCTCTGTTTACTTCTACAAATCCAAAGAGATGAAACAGATGTGACAGATTTATCATATT 2100
Qy 2032 GATCAAGTATCCAATTTAGTGGCGTGTATTCGGATGAATTTCTGTTAGATGAAAGAGA 2091
Db 2101 GACCAAGTGTCCAAATATGTTGGCGCATGTTTATCAGATGAATTTTGTGATGAGAAGCGA 2160
Qy 2092 GAATTTACTTGAGAAAGTGAATATATGCAAAACGACTCAGTGATGAAAGAACTTACTCCAA 2151
Db 2161 GAATTTATTTGAGAAAGTGAATATATGCGAAGCGACTCAGTGATGAAAGAACTTACTCCAA 2220
Qy 2152 GATCCAAACTTTCACATCCATCAATAAGCAACAGACTTTCTATCTACTACTGATGCAATTCG 2211
Db 2221 GATCCAAACTTTCACATTCAGTGGGCAATTAAGTTTTCGATCCATCGATGCAATCA 2280
Qy 2212 AATTTTCATCTATCCATGAACATCTGAAACATGATGGTGGGAGAGTGAAGCAATTACA 2271
Db 2281 AACTTCCCTCTATTAATGAGCTATCTGAACATGGATGGTGGGAGAGTGGCAATTTTACC 2340
Qy 2272 ATCCAGGAAGGAAATGAGCTATTTAAAGAGAAATACGTCACACTACCGGGACTTTTAAT 2331
Db 2341 ATTACGGAAGGAAATGAGCTATTTAAAGAGAAATACGTCACACTACCGGGACTTTTAAT 2400
Qy 2332 GAGTGTATTCGACGCTATTTATATCAAAAAATAGGAGTGGGAATTTAAAGCTTATACT 2391
Db 2401 GAGTGTATTCCAAAATTTATATCAAAAAATAGGAGTGGGAATTTAAAGCTTATACG 2460
Qy 2392 CGCTACCAATTAAGAGGTATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGT 2451
Db 2461 CGCTATCAATTAAGAGGTATATTTGAAGATAGTCAAGATCTAGAGATTTTATTTAATTCGT 2520

Qy 2452 TATAATGCGAAACATGAAACATTTGATGTTTCCAGGTACCGAGTCCGTATGCGCGCTTTCA 2511
Db 2521 TACAATGCAAGCATGAAACATTTGATGTTTCCAGGTACCGATTCCTATGCGCGCTTTCA 2580
Qy 2512 GTTGAAGCCCCAATCGGAGGTGCGGAGAACCGAATCGATCGGCACCAATTTTGAATGG 2571
Db 2581 GTTGAAGCCCCAATCGGAGGTGCGGAGAACCAATCGATCGGCACCAATTTTGAATGG 2640
Qy 2572 AATCTGATCTAGATTTGTTTCTGAGAGATGAGAGAAATTTGCGCATCATTTCCCATCAT 2631
Db 2641 AATCTGATCTAGATTTGTTTCTGAGAGATGAGAGAAATTTGCGCATCATTTCCCATCAT 2700
Qy 2632 TTCTCTTTGGATATTTGATTTTGGATGCAACACTTTGCAATGCAATCTAGGCGTGTGGTG 2691
Db 2701 TTCTCTTTGGATATTTGATTTTGGGTGCAACACTTTGCAATGCAATCTAGGCGTGTGGTG 2760
Qy 2692 GTATTTCAAGATTTAGACCGCAGGAGGTTCATGCAAGCTAGGCAATCTGGAATTTATTGAA 2751
Db 2761 GTATTTCAAGATTTAGACCGCAGGAGGTTCATGCAAGATTTAGGAAATCTGGAATTTATTGAA 2820
Qy 2752 GAGAAACCATTTATAGGAGAGCACTGTCTGTGAAAGAGAGAGAGAGAGAGAGAGAGAG 2811
Db 2821 GAGAAACCATTTATTTGAGAGAGCACTGTCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 2880
Qy 2812 GACAAACGTGAAAACTCAATTTGAAAAAATAAAGAGATATATACAGAGGCAAAAGAGCT 2871
Db 2881 GACAAACGGGAAAACTCAATTTGAAAAAATAAAGAGATATATACAGAGGCAAAAGAGCT 2940
Qy 2872 GTGATGCTTTATTTGATGATTTCTCAATATATATAGATTTCAAGCGGATACAAACATTGGC 2931
Db 2941 GTGATGCTTTATTTGATGATTTCTCAATATATATAGATTTCAAGCGGATACAAACATTGGC 3000
Qy 2932 ATGATTTCAATGCGCAGATATAACTTTGTCATGCAATTCGAGAGGCTTATCTGTCAGAAATTA 2991
Db 3001 ATGATTTCAATGCGCAGATATAACTTTGTCATGCAATTCGAGAGGCTTATCTTTAGAAATTA 3060
Qy 2992 TCTGTTATCCCGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAGGTCGCAATTTACT 3051
Db 3061 CTTGTTATCCCGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAGGTCGCAATTTACT 3120
Qy 3052 GCAATCTCCCTATACGATGCGAGAAATGTCGTTAAAAATGCTGATTTTAAATAGGATTA 3111
Db 3121 GCAATGCTTTATACGATGCGAGAAATGTCGTTAAAAATGCTGATTTTAAATAGGATTA 3180
Qy 3112 GCATGCTGGAATGTAAAGGCGATGATAGTGTACACAGAGCCATCACGTTCTGTCCTT 3171
Db 3181 ACATGTTGGAATGTAAAGGCGATGATAGTGTACACAGAGCCATCATCTGTTGACCTT 3240
Qy 3172 GTTATCCAGAAATCGGAGAGCAAGTGTCAACAGCAGTTCCGCTCTGTCGCGGGCGTGGC 3231
Db 3241 GTTATCCAGAAATCGGAGAGCAAGTGTCAACAGCAGTTCCGCTCTGTCGCGGGCGTGGC 3300
Qy 3232 TATATCTCTGCTGTCAACGCGTACAAAGAGGATATGAGAGGTTGTGTAAACGATCCAT 3291
Db 3301 TATATCTCTGCTGTCAACGCGTACAAAGAGGATATGAGAGGTTGTGTAAACGATCCAT 3360
Qy 3292 GAATTCGAGAACATACAGAGCAACTAAATTTTAAACCTGCTGAGAGAGAGAGAGTGTAT 3351
Db 3361 GAATTCGAGAACATACAGAGCAACTAAATTTTAAACCTGCTGAGAGAGAGAGAGTGTAT 3420
Qy 3352 CCAACGATACAGGAAGTGTAAATGATTTATCTGCAACCAAGGTACAGCAGTATGTAAT 3411
Db 3421 CCAACGATACAGGAAGTGTAAATGATTTATCTGCAACCAAGGTACAGCAGTATGTAAT 3480
Qy 3412 TCCGTTAATGCTGGATATGAGGATGCAATAGAGTTGATCTACTACAGCATCTGTTAAATPAC 3471
Db 3481 TCCGTTAATGCTGGATATGAGGATGCAATAGAGTTGATCTACTACAGCATCTGTTAAATPAC 3540
Qy 3472 AAACCGACTTATGAAGAGAAACGATATACAGATGTACAGAGATATCATTTGTAATAT 3531
Db 3541 AAACCGACTTATGAAGAGAAACGATATACAGATGTACAGATGTACAGAGATATCATTTGTAATAT 3600

Db 3361 GAAATCGAGAAACAATA CAGACGAACTAAAAATTTAAAAAAGAGGAAAGTGTAT 3420
Qy CCAACGGATACAGGAACGTGTAATGATTATCTGCACACCAAGGTACAGCAGTATGTAAT 3411
Db CCAACGGATACAGGAACGTGTAATGATTATCTGCACACCAAGGTACAGCAGTATGTAAT 3480
Qy TCCCGTAAATGCTGGATATGAGGATGCAATATGAAGTTGATACCTACAGCATCTGTTAAATAC 3471
Db TCCCGTAAATGCTGGATATGAGGATGCAATATGAAGTTGATACCTACAGCATCTGTTAAATAC 3540
Qy AAACCGCATTTGAAGAGAAACGTTATACAGATGACAGAGATATCATTTGTAATAT 3531
Db AAACCGCATTTGAAGAGAAACGTTATACAGATGACAGAGATATCATTTGTAATAT 3600
Qy GACAGAGGTTATGTAATATCCACCACTACCACTGTTTATATGACAAAGAAATTAAGAA 3591
Db GACAGAGGTTATGTAATATCCACCACTACCACTGTTTATATGACAAAGAAATTAAGAA 3660
Qy TACTTCCAGAAACCGATAGGATGATGAGATTGAGATTGAGATTGAGATTGAGATTGAGATT 3651
Db TACTTCCAGAAACCGATAGGATGATGAGATTGAGATTGAGATTGAGATTGAGATTGAGATT 3720
Qy GTACAGAGCGTGGAAATTTACTCTTATGAGGAAATAG 3687
Db GTAGATAGCGTGGAACTACTCTCATGGAAGATAG 3756

RESULT 10
US-08-881-340-3
; Sequence 3, Application US/08881340
; Patent No. 5942658
; GENERAL INFORMATION:
; APPLICANT: Donovan, William P.
; APPLICANT: Tan, Yiping
; APPLICANT: Jan, Christine S.
; APPLICANT: Gonzalez Jr., Jose M.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS cryET4 AND cryET5
; TITLE OF INVENTION: TOXIN GENES AND PROTEINS TOXIC TO LEPIDOPTERAN INSECTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Panitch Schwarze Jacobs & Nadel c/o A.S.
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881.340
; FILING DATE: 24-JUN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/100,709
; FILING DATE: 29-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Egoif, Christopher
; REGISTRATION NUMBER: 27633
; REFERENCE/DOCKET NUMBER: 7205-49
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-757-1590
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3934 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; FEATURE:

; NAME/KEY: CDS
; LOCATION: 67..3756
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2253..2272
US-08-881-340-3

Query Match 72.9%; Score 2687.2; DB 2; Length 3934;
Best Local Similarity 83.5%; Pred. No. 0;
Matches 3088; Conservative 0; Mismatches 593; Indels 15; Gaps 3;

Qy 1 TTGACTTCAAATAGGAAAAATGAGAAATGAAATATATAAATGCTTTATCGATTCCAGCTGTA 60
Db 67 TTGACTTCAAATAGGAAAAATGAGAAATGAAATATATAAATGCTTTATCGATTCCAGCTGTA 126
Qy 61 TCGAATCATTTCCACACAAATGGAATCTATCACAGATGCTCGTATTTGAGGATTTCTTCTGT 120
Db 127 TCGAATCTTTCCACGCAAAATGAATCTATCACAGATGCTCGTATTTGAGGATTTCTTCTGT 186
Qy 121 ATAGCCGAGGGGAATATATCAATCCACTTTGTTAGCGCATCAACAGTCCAAACGGGTATT 180
Db 187 GTAGCCGAGGGGAATATATCAATCCACTTTGTTAGCGCATCAACAGTCCAAACGGGTATA 246
Qy 181 AACATAGCTGGTAGAATACCTAGGTGTTATAGGCGTACCGTTTGCCTGGACAAAATAGCTAGT 240
Db 247 AACATAGCTGGTAGAATATTTGGGCGTATTAGGTGTTGCGGTTTGGTGACAACTAGCTAGT 306
Qy 241 TTTTATAGTTTCTTGTGTTGTAATATGCGCCGCGGCGAGAGATCAGTGGGAAATTTTC 300
Db 307 TTTTATAGTTTCTTGTGTTGGAATATGAGCGCTAGTGGCAGAGATCCATGGGAAATTTTC 366
Qy 301 CTAGAACATGTCGAAACAACTTATAAAATCAACAAATAACAGAAAATGCTAGGAATACGGCA 360
Db 367 CTGGAACATGTAGAACAACTTATAAGACAACTACAGAAATACTAGGAATACGGCT 426
Qy 361 CTGCTCGATTACAGGTTTAGGAGATTCCTTTAGAGCCTATCAAGTCACTTGAAGAT 420
Db 427 ATTGCTCGATTAGAGGTTTAGGAGAGGCTATAGATCTTACCAGCAGGCTCTTGAAGT 486
Qy 421 TGGCTAGAAACCGTGTATGATGCAAGAACGAGAGTGTCTTTTATATCCCAATATATAGCC 480
Db 487 TGGTTAGATTAACCGAAATGATGCAAGATCAAGAGCAATATCTTTGAGCGCTATGTTGCT 546
Qy 481 TTAGAACTTCGATTTCTTAATGCGATGCCCTTTTCGCAATTAGAAACCAAGAGTTTCCA 540
Db 547 TTAGAACTTCGATTTCTTAATGCGATGCCCTTTTCGCAATTAGAAACCAAGAGTTTCCA 606
Qy 541 TTATTAATGTTATGCTCAAGCTGCAAAATTTACCTATTTATTTGAGAGATGCCTCT 600
Db 607 TTATTAATGTTATGCTCAAGCTGCAAAATTTACCTATTTATTTGAGAGATGCCTCT 666
Qy 601 CTTTTTGGTAGTAATTTGGCTTACATCGCAGGAATTTCAAGTTTATTTATGAGCGCAA 660
Db 667 CTTTTTGGTAGTAATTTGGGGAATTTGGCAATCTTCGATGTTAACCAATATTTACCAAGAACAA 726
Qy 661 GTGGAACAAACGAGAGATTTATCCGACTATTGGCTAGAAATGCTATATACAGGTTCTAAAT 720
Db 727 ATCAGATATACAGAGGAATTTCTAACCAATTTGGTACAAATGTTATATACAGGCTAAAT 786
Qy 721 AGCTTAGAGGGAACAATGCCGAAAGTTGGGTGGCTTATTAATCAATTTCCGTAGAGATCTA 780
Db 787 AACCTTAAGAGGGAACAATGCTGAAAGTTGGTGGGTATTAATCAATTTCCGTAGAGACTA 846
Qy 781 ACCTTAGGGGTATTTAGATCTAGTGGCACTATTTCCCAAGCTATGACACTCGCACCTATCCCA 840
Db 847 ACCTTAGGGGTATTTAGATCTAGTGGCACTATTTCCCAAGCTATGATATCTCGCACCTATCCCA 906
Qy 841 ATAAATACGAGTGTCTCAGTTTAAACAAAGGAAGTTTATACAGACGCAATTTGGAGCAACAGGG 900
Db 907 ATCAATACGAGTGTCTCAGTTTAAACAGAGAAATTTATACAGATCCAAATTTGGAGAACAAAT 966
Qy 901 GTAAAT-----ATGGCAAGTATGAAATGGTATTAATTAATGACCTTGTGTTTTCGCT 954

Db	967	GCACCTTCAGGATTTGCAAGTACGAATTTGGTTTAAATAATAATGACCACTGCTTTCTGCC	1026
Qy	955	ATAGAGACTGGGTTATCCGAAGCCCGCATCTACTTGAATTTTCTAGAAACAATTTACAAT	1014
Db	1027	ATAGAGGTCGCAATTTTCAGGCTCCGATCTACTTGAATTTTCCAGAACAACTTACAAT	1086
Qy	1015	TTTAGCACTTCATCAGCATGAGTGTCTACTAGGCATATGACTTACTTGGCGGGGACACACA	1074
Db	1087	TACAGTGCATCAAGCCGTTGGAGTAGCACTCAACATATGAAATTTATGGGTGGGACATAGG	1146
Qy	1075	ATTCAATCTCGGCCAATAGGAGGCGGATTAATACTCTCAACGCGATGGGTCTACCA--AT	1131
Db	1147	CTTAATCTCCGCCCAATAGGAGGACATTAATACTCTCAACAGAGCTTACTAATAAT	1206
Qy	1132	ACTTCTATTAATCTGTGAAGATTATCAATCTTCTCTCGAGACGTATATTTGGACTGAATCA	1191
Db	1207	ACTTCAATTAATCTGTGAACATTAACAGTTAGCTCTCGAGAGCTTTATAGAACAGATCA	1266
Qy	1192	TATGAGGAGTGTCTTATGGGAAATTTACCTTGNACCTATTCATGGTGTCCCTACTGTT	1251
Db	1267	AATGACGGGACAAATAT-----ACTATTTACTCTCCTGTGAATGGAGTACCTTGGGCT	1320
Qy	1252	AGATTTAATTTTAGGAACCTCAGAACTATTTTGAAGAGGTACTGTCTAACTATAGTCAA	1311
Db	1321	AGATTTAATTTTAAACCTCAGAAATATTTATGAAGAGGCGCACTACCTACAGTCAA	1380
Qy	1312	CCCTATGAGTCACTGGGCTTCAATTAAGAAATTCAGAAACTGAATTAACCAACCGAAACA	1371
Db	1381	CGGTATCAGGGAGTTGGGATTTCAATTTATTTGATTTCAGAAACTGAATTAACCAACCGAAACA	1440
Qy	1372	ACAGAACGACCAATTAATGAATCATATAGTCATAGGTTATCTCATAAGGGCTCAATTTCA	1431
Db	1441	ACAGAACGACCAATTAATGAATCATATAGTCATAGTATTAATCTCATATAGGACTAATCATA	1500
Qy	1432	CAATCTAGGTCATGTACCACTATATTTCTGGAGCAGCCGTAGTCAGATCGTACAAAT	1491
Db	1501	GGAAACCTTTGAGAGCACCCAGTCTATTTCTGGAGCAGCATCTGATGCGACATCGTACGAAT	1560
Qy	1492	ACCATTAGTTCAGATPAGCATAAACAAATACCATTTGGTAAATCATTTCAACCTTAATTTCA	1551
Db	1561	ACGATTGACCAATAGAATTTACAAATACCAATTTGGTAAAGCACTGNACTTTCATTTCA	1620
Qy	1552	GGTACCTCTGTAGTCAGTGGCCAGGATTTACAGAGGGGATATAATCCGAACCTAAAGTT	1611
Db	1621	GGTGTACTGTGTGGAGGCGCCAGGATTTACAGTGGGGATATCCCTTCGTAGAACAAAT	1680
Qy	1612	AATGGTACTGTACTAAGTATGGGCTTAAATTTTAAATAATCATATACATACACGCGTATCGC	1671
Db	1681	ACGGGTACATTTGGAGATATACGATTAATAATTAATGTGCCATTTATCCCAAGATATCGC	1740
Qy	1672	GTGAGAGTTCTGTATGCTGCTTCTCAACAAATGGTCTCTGAGGGTAACTGTGCGAGGGAGT	1731
Db	1741	GTAAGGATTCGTTATGCTTCTTACTACAGATTTACAAATTTTTCACAGAAATTAATGGAACC	1800
Qy	1732	ACTACTTTTGTCAAGGATTCCTTAGTACTATGAGTGCAATGAGTCTTTGACATCTCAA	1791
Db	1801	ACTGTTAATATTTGTAATTTCTCAAGAACTATGAATAGGCGGGATATTTAGAAATATAGA	1860
Qy	1792	TCATTTAGATTTGAGAAATTTCTCTGTAGGTATTAAGTGCATCTGGCAGTCAAACTGCTGGA	1851
Db	1861	AGTTTTAGAACTGCGAGGATTTAGTACTCTCTTTTAAATTTTAAATGCCCCAAAGCACATTC	1920
Qy	1852	ATAAGTATAAGTAAATGACAGGTAGACAAACGTTTCACTTTGATATAAAATGAAATTCAT	1911
Db	1921	ACATTTGGGTGCTCAGAGTTTTCAAATCAGGAAGTTTATATAGATAGAGTGCATTTGTT	1980
Qy	1912	CCAATTTACTGCAACCTTCGAAGCAGAAATPACGATTTAGAAAGGCGGCAAGAGCGGTGAAT	1971
Db	1981	CCAGCAGAGGTAACATTTTGGAGCAGAAATATGATTTAGAAAGCAGCACAAAGCGGTGAAT	2040
Qy	1972	GCTCTGTTTACTAATACGAATCCGAAGAGATGAAACAGATGTGACAGATTTATCATATT	2031
Db	2041	GCTCTGTTTACTTCTACAAATCCGAAGAGATTTGAAACAGATGTGACAGATTTATCATATT	2100
Qy	2032	GATCAAGTATCCAAATTTAGTGGCGTGTATTCGGATGAATTTCTGCTTAGATGAAGAAGA	2091
Db	2101	GACCAAGTGTCCAATATGGTGGCATGTTTACAGATGAATTTGCTTGGATGAGAAGCGA	2160
Qy	2092	GAATTACTTTGAGAAAGTCAANAATGCCGAACGACTCAGTGATGAAGAAGAACTTACTCCAA	2151
Db	2161	GAATTAATTTGAGAAAGTGAANAATGCCGAAGCGACTCAGTGATGAAGAAGAACTTACTCCAA	2220
Qy	2152	GATCCAAACTTCAATCCATCAATAAGCAACACGACTTCATATCTTACTTAATGAGCAATCG	2211
Db	2221	GATCCAAACTTCAATTCATCAGTGGGCAATTAAGTTTTCGCATCCATCGATGGCAATCA	2280
Qy	2212	AAATTCACTATCTCCATGAACAAATCGAATCGATGGTGGGGAAGTGAGAACATTTACA	2271
Db	2281	AACTTCCCTCTTAATAATGAGCTATCTGAACATGATGGTGGGGAAGTGCGAATGTTACC	2340
Qy	2272	ATCCAGGAAGGAATGAGTATTTAAAGAGAAATTTACGTCACTACACACCGGGTACTTTTAA	2331
Db	2341	ATTCAGGAAGGAATGAGTATTTAAAGAGAAATTTACGTCACTACACACCGGGTACTTTTAA	2400
Qy	2332	GAGTGTATCCGACGTATTTATATCAAAAAATAGCAGAGTCGGAATTAAGAGCTTATACT	2391
Db	2401	GAGTGTATCCAAATTTATATCAAAAAATAGCAGAGTCGGAATTAAGAGCTTATAGG	2460
Qy	2392	CGTACCAATTAAGAGGGTATTTGAAGATAGTCAAGATTTTAGAGATATATTTGATTCGT	2451
Db	2461	CGCTATCAATTTAAGAGGGTATTTGAAGATAGTCAAGATCTAGAGATTTATTTAATTCGT	2520
Qy	2452	TATTAATCGGAACATGAACATTTGGATGTTTCAGGTACCGAGTCGATATGCGCGCTTTCA	2511
Db	2521	TACAATCGGAACATGAACATTTGGATGTTTCAGGTACCGATTTCCCTATGCGCGCTTTCA	2580
Qy	2512	GTTCAGAGCCCAATCGGAAGGTGGGAGAACCGAATCGATCGGCACACATTTTGAATGG	2571
Db	2581	GTTCAGAGCCCAATCGGAAGGTGGGAGAACCGAATCGATCGGCACACATTTTGAATGG	2640
Qy	2572	AATCCTGATCTAGATTTGCTCTGCGAGAGATGGAGAAAAATGTGGCGATCATTTCCCATCAT	2631
Db	2641	AATCCTGATCTAGATTTGCTCTGCGAGAGATGGAGAAAAATGTGGCGATCATTTCCCATCAT	2700
Qy	2632	TTCTCTTTGGATATTTGATTTGATTTGACAGACTTTGATGAGAACTTAGGGGTGGGTG	2691
Db	2701	TTCACTTTGGATATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG	2760
Qy	2692	GTATTTCAAGATTAAGACGCGAGAGGTCTGCAAGACTTAGGGAATCTGGAATTTTATTCGA	2751
Db	2761	GTATTTCAAGATTAAGACGCGAGAGGTCTGCAAGACTTAGGGAATCTGGAATTTTATTCGA	2820
Qy	2752	GAGAAACCATTTATAGGAGAGCACTGCTCTGCTGTGAAGAGAGAGAGAAAAATGGAGA	2811
Db	2821	GAGAAACCATTTATTTGAGAGAGCACTGCTCTGCTGTGAAGAGAGAGAGAAAAATGGAGA	2880
Qy	2812	GACAAACGTGAAAAATCTACAAATTTGGAACAAAAACGAGTATATACAGAGGCAAAAGAGCT	2871
Db	2881	GACAAACGTGAAAAATCTACAAATTTGGAACAAAAACGAGTATATACAGAGGCAAAAGAGCT	2940
Qy	2872	GTGGATGCTTTATTTGATGATTTCTCAATATATAGATTTACAGCGGATACAAACATTTGGC	2931
Db	2941	GTGGATGCTTTATTTGATGATTTCTCAATATATAGATTTACAGCGGATACAAACATTTGGC	3000
Qy	2932	ATGATTCATTCGCGCAGATAAATTTGTTTCACTGGAATTCGAGAGGCTTATCTCTCAGAAATTA	2991
Db	3001	ATGATTCATTCGCGCAGATAAATTTGTTTCACTGGAATTCGAGAGGCTTATCTCTCAGAAATTA	3060
Qy	2992	TCTGTTATCCGGGTGTAATTTGCGGAAATTTTGAAGAAATTTGAGAGGTCGCAATTCACAT	3051
Db	3061	CCTGTTATCCGGGTGTAATTTGCGGAAATTTTGAAGAAATTTGAGAGGTCGCAATTCACAT	3120
Qy	3052	GCAATCTCCCTATACGATGCGAGAAATGCTGTTAAAAATGCTGATTTTAAATGGAATTA	3111
Db	3121	GCAATGCTCTTATACGATGCGAGAAATGCTGTTAAAAATGCTGATTTTAAATGGAATTA	3180

```

Qy 3112 GCATGCGGAATGTAAGGCGCATGTAGATGTACAAAGAGCCATACCGTCTGTCTT 3171
Db 3181 ACATGTTGGAATGTAAGGCGCATGTAGATGTACAAAGAGCCATACCGTCTGTCTT 3240
Qy 3172 GTTATCCAGATGGGAGAGAGAGTGTCAAGAGTTCGCTCTCCGGGCGTGGC 3231
Db 3241 GTTATCCAGATGGGAGAGAGAGTGTCAAGAGTTCGCTCTCCGGGCGTGGC 3300
Qy 3232 TATATCTCTCGTGTCAAGAGTGTCAAGAGGATGTAGAGGCTGTGTAAACCATCAT 3291
Db 3301 TATATCTCTCGTGTCAAGAGTGTCAAGAGGATGTAGAGGCTGTGTAAACCATCAT 3360
Qy 3292 GAAATCCAGAACATACAGACGAATATAAATTTAAACTGTGTGAAGAGGAGTGTAT 3351
Db 3361 GAAATCCAGAACATACAGACGAATATAAATTTAAACTGTGTGAAGAGGAGTGTAT 3420
Qy 3352 CCAACGATACAGAAAGTGTAAATGATTATATCTCCACCAAGGTACAGCAGTGTAAAT 3411
Db 3421 CCAACGATACAGAAAGTGTAAATGATTATATCTCCACCAAGGTACAGCAGTGTAAAT 3480
Qy 3412 TCCCGTAAATGCTGATATGAGATGCTATATGAAGTGTGATCTACAGCATCTGTAAATAC 3471
Db 3481 TCCCGTAAATGCTGATATGAGATGCTATATGAAGTGTGATCTACAGCATCTGTAAATAC 3540
Qy 3472 AAACCGACTTATGAAGAGAAACGATACAGATGTACGAAGAGATAATCAATGTGAATAT 3531
Db 3541 AAACCGACTTATGAAGAGAAACGATACAGATGTACGAAGAGATAATCAATGTGAATAT 3600
Qy 3532 GACAGAGGATGTGAATATTCACCACTACCACTGCTGTATATGACAAAGAAATAGAA 3591
Db 3601 GACAGAGGATGTGAATATTCACCACTACCACTGCTGTATATGACAAAGAAATAGAA 3660
Qy 3592 TACTTCCAGAAACCGATAGGTATGATGAGATTCGAGAAACGAGGAGGAGTGTATT 3651
Db 3661 TACTTCCAGAAACCGATAGGTATGATGAGATTCGAGAAACGAGGAGGAGTGTATT 3720
Qy 3652 GTACAGAGCGTGAATATCTCTTTATGAGGAATAG 3687
Db 3721 GTACAGAGCGTGAATATCTCTTTATGAGGAATAG 3756

```

```

RESULT 11
US-08-040-751-4
; Sequence 4, Application US/08040751
; Patent No. 5407825
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M
; APPLICANT: Sick, August J
; TITLE OF INVENTION: No. 5407825el Bacillus thuringiensis isolates
; TITLE OF INVENTION: active against Lepidopteran Pests and Genes Encoding No. 54078
; TITLE OF INVENTION: Lepidopteran-active Toxins
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID R. SALIWANCHIK
; STREET: 2421 N.W. 41st STREET, SUITE A-1
; CITY: GAINESVILLE
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/040,751
; FILING DATE: 19930329
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SALIWANCHIK, DAVID R.
; REFERENCE/DOCKET NUMBER: MA39.C1.D3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100

```

```

; TELEFAX: 904-372-5800
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3522 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: aizawai
; INDIVIDUAL ISOLATE: P881A2
; IMMEDIATE SOURCE:
; LIBRARY: Lambdaem - 11 (tm) Library of August Sick
; CLONE: 81A2
; US-08-040-751-4
;
; Query Match 48.6%; Score 1793.4; DB 1; Length 3522;
; Best Local Similarity 71.9%; Pred. No. 0;
; Matches 2497; Conservative 0; Mismatches 861; Indels 117; Gaps 7;
;
; Qy 228 ACAATAGCTAGTGTATTTATAGTGTCTTCTGTGTGTAATATGCGCCCGCGCAGAGATCA 287
; Db 147 ACTAGGGGATTTTATATCTTGGCTTGTGTATGTAATATGCGGGGCTATAGTCTCTTCACA 206
;
; Qy 288 GTGGGAAATTTCTTAGAACATGTGNAACATTTATAAATCAACAATACAGAAATGC 347
; Db 207 ATGGGATATATTTTAGAGCAAAATGAGCTTATTGATCGGCCAAAGAAATAGAGGAATTCG 266
;
; Qy 348 TAGGAATACGGCACTTGTCTCGATTACAAGGTTTAGGAGATTCCTTTAGAGCCTATCAACA 407
; Db 267 TAGGAATACGGCAATTTCTAGATTACAGGCTTAAGCAATCTTTACCGAATTTACACAAA 326
;
; Qy 408 GTCATTTGAAGATGCGCTAGAAAACCGTGAATGATGAAAGAACGAGAGTGTCTTTTATAC 467
; Db 327 TGCCTTTTAAAAAATCTGGGAGTAGATCCTACTAATCCAGCATTAAGAGAGAGATGCGTAT 386
;
; Qy 468 CCAATATAGCTTAGAACTTGTCTTCTTAATGGATCGCGCTTTTCCCAATTAGAAA 527
; Db 387 TCAATTTAATGATGATGAAAGAGTGTCTTTTCAACAGCTATTCTCTTTTTCAGTTCAGG 446
;
; Qy 528 CCAAGAGTTTCCATTTAATGATGATGCTCAAGCTGCAAAATTTACACCTATTATTATT 587
; Db 447 TTATGAAATTCCTTTTATCAGTATATGTTCAAGCTGCAAAATTTATACATTAATCGGTTT 506
;
; Qy 588 GAGAGATGCTCTCTTTTGGTGTAGTAATTTGGGCTTACATCGCAGGAAATTCACGTTA 647
; Db 507 GAGAGATGTTTCAAGTGTTCGACAAACGTTGGGATTTGATGTAGCAACAATCAATAGTCG 566
;
; Qy 648 TTATGAGCGCAAGTGGAAACAAACGAGAGATTTCCGACTATTTCGTAGATGCGTATAA 707
; Db 567 TTATATGATTTAACTAGGCTTTATGCGGAATATATCTGATGCTGCTGCTGCTGATAA 626
;
; Qy 708 TACAGGTCTAAATAGCTTTGAGAGGCAAAATGCGG---CAAGTTGGGTGGCTTATAATCA 764
; Db 627 TACGGGTTAAATCGTTTACCACTAATGAAGGGTACGAGGATGGCAAGATTTAATAG 686
;
; Qy 765 ATTCCGCTAGAGATCTAACGTTAGGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGA 824
; Db 687 GTTTAGAAAGAGAGTTAAACAATATCAGTATTAGATATTATTTCTTTTCCAAAATTCAGA 746
;
; Qy 825 CACTCGCACTTATCCCAATTAATAGAGTGTCTCAGTTAAACAGGGAAGTATTATACAGACG 884
; Db 747 TTCTAGATTTATCCCAATTCAGCAATCTATCAATTAACCGGGAAGTATTATACAGATCC 806
;
; Qy 885 AATTGGAGCAACAGGGGTAAATATGCAAGTATGCAATTTGGTATAATTAATGACACCTTC 944
; Db 807 -----GGTAATTAATATACTGATTTATAGAGTTACCCCAAG 842
;
; Qy 945 GTTTTCGCTATAGAGACTCGGGTTATCCGAAGCCCGCACTCTACTGATTTTCTAGAAC 1004

```


QY 3162 TTCTGTCTTCTTTATCCAGAAATGGGAGCAGAAAGTGTCAACAGCAGTTCGCGTCTGCC 3221
Db 2988 TTGGTCTCTTGTTCCTCGGAATGGGAGCAGAAAGTGTCAACAAATTCGTGTCTGCC 3047
QY 3222 GGGCGTGGCTATATCTCTCGTGTCAAGCGTACAAAGAGGATATGGAGGGTGTGT 3281
Db 3048 GGGCGTGGCTATATCTCTCGTGTCAAGCGTACAAAGAGGATATGGAGGGTGTGT 3107
QY 3282 AACGATCCATGAATCCAGACAAATACAGACGAACTAAATTTAAACCTCTGAAGAAGA 3341
Db 3108 AACCATCATGAGATCGAGACAAATACAGACGAACTAAATTTAAACCTCTGAAGAAGA 3167
QY 3342 GGAAGTGTATCCAAAGGATACAGAAAGCTGTAAATGATTATCTCCACCAAGGTACAGC 3401
Db 3168 GGAAGTGTATCCAAAGGATACAGAAAGCTGTAAATGATTATCTCCACCAAGGTACAGC 3227
QY 3402 AG-----TATGTAAATCCCGTAATGCTGGATATGAGATGCATATGAAGTTGA 3449
Db 3228 AGGATCCACAGATTCATGTAAATCCCGTAATGCTGGATATGAGATGCATATGAAGTTGA 3287
QY 3450 TACTACAGCATCTGTAAATCAACACGACTTATGAAGAGAAACGATATACAGATGTACG 3509
Db 3288 TACTACAGCATCTGTAAATCAACACGACTTATGAAGAGAAAGGTATACAGATGTACG 3347
QY 3510 AAGAGATAATCATTTGTGAATATGACAGAGGGTATGTGAATTTATCCACCACTTACCAGCTGG 3569
Db 3348 AGGAGATAATCATTTGTGAATATGACAGAGGGTATGTGAATTTATCCGACAGTACCAGCTGG 3407
QY 3570 TTATATGACAAAGAAATAGAAATCTCCAGAAACCGATAGATGTGAATGTGATGG 3629
Db 3408 TTATGTGACAAAGAAATAGAGTACTTCCAGAAACCGATAGATGTGAATGTGATGG 3467
QY 3630 AGAAACGGAAGGAAGTTATTTAGACAGCGTGAATTAATCTCTTATGAGGAA 3684
Db 3468 AGAAACGGAAGGAAGTTATTTAGACAAATGTGCAATTAATCTCTTATGAGGAA 3522

RESULT 12

US-08-291-368-1
; Sequence 1, Application US/08291368
; Patent No. 5686069
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Sick, August J.
; TITLE OF INVENTION: No. 5686069el Bacillus thuringiensis Isolates
; TITLE OF INVENTION: Active Against Lepidopteran Pests
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESS: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/291.368
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/597,607
; FILING DATE: 15-OCT-90
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA50.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904)375-8100
; TELEFAX: (904)372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3522 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: aizawai
; INDIVIDUAL ISOLATE: P881A2
; IMMEDIATE SOURCE:
; LIBRARY: Lambdegem - 11 (cm) Library of August Sick
; CLONE: 81A2
; US-08-291-368-1
Query Match 48.6%; Score 1793.4; DB 1; Length 3522;
Best Local Similarity 71.9%; Pred. No. 0;
Matches 2497; Conservative 0; Mismatches 861; Indels 117; Gaps 7;
QY 228 ACAAATAGCTAGTGTCTTATAGTTTCTTGTGTGAAATATGCGCCCGCGAGAGATCA 287
Db 147 ACTAGGGGATTTTATATCTTGGCTTGTGTGATGTAATATGGGGCTATAGTCTCTTCA 206
QY 288 GTGGGAATTTTCTTAGAACATGTCGAACTTATAAATCAACAATACAGAAATGC 347
Db 207 ATGGGATATATTTTAGAGCAAAATGAGCTTATGTCGGCCAAAGATAGAGGAATTCG 266
QY 348 TAGCAATACCGCACTTCTCGATTACAAGGTTTAGGAGATTCTTTTAGAGCCTATCAACA 407
Db 267 TAGAATCAGCAATTTCTAGATTACAAGGCTTAAGCAATCTTTACCGAATTTACACAA 326
QY 408 GTCATTGAAGATGGCTAGAAAACCGTGATGATGCAAGAACGAGAGAGTGTCTTTATAC 467
Db 327 TGCTTTTAAAAAATCTGGGAGTAGATCCTACTAATCCAGCAATTAAGAGAGAGATCG 386
QY 468 CCAATATAGCTTAGAATCTGATTTCTTATCGATGCGCTTTTCGCAATTAGAAA 527
Db 387 TCAATTTAATGATGACAGACAGTGTCTTACAAGCTATTCCTCTTTTTCAGTTCAAGG 446
QY 528 CCAAGAGTTCCTAATTTAATGTTATGCTCAAGCTGCAAAATTTACACCTATTTATT 587
Db 447 TTATGAATTCCTCTTTTATCAGTATATGTTCAAGCTGCAAAATTTACATTTACGGT 506
QY 588 GAGAGATGCTCTCTTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATCAACGTTA 647
Db 507 GAGAGATGTTTCAAGTGTTCGACAAACGTTGGGATTTGATGTAGCAACAATCAATAG 566
QY 648 TTATGAGCGCAAGTGGAAACAAACGAGAGATTATTCGACTATTCGTAATGCGTATAA 707
Db 567 TTATAATGATTTAATAGGCTTATTTGGCGAATATATCTGATTTATGCTGCTGTTATA 626
QY 708 TACAGGTCTAAATAGCTTTGAGAGGCAAAATCGCG---CAAGTTGGGTGCTTTATAATCA 764
Db 627 TAGGGGTTAAATCGTTTACCAAGTAAAGGGGTACGAGGATGGCAAGATTTAATAG 686
QY 765 ATTCGCTAGAGATCTAACGTTAGGGGTATTTAGATCTAGTGGCACTATTTCCAGAGTATGA 824
Db 687 GTTTAGAAAGAGAGTTAACCAATATACATATATTTATTTCTTTTTCCTTTTCAAAAT 746
QY 825 CACTCCGCACTTATCCCAATTAATACGAGTGTCTAGTTTAAACAGGGAAGTTTATACAGACG 884
Db 747 TTCTAGATTTATCCCAATTCGCAATCTATCAATTAACCGGGAAGTATATACAGATCC 806
QY 885 AATTGGAGCAACAGGGGTAAATATGCAAGTATGCAATTTGGTATAATTAATGACACCTTC 944
Db 807 -----GGTAAATTAATAACTGATTATAGAGTATACCCCAAG 842
QY 945 GTTTTCGCTATAGAGACTCGGTTATCCGAAGCCCGCATCTACTTGATTTTCTAGAACA 1004

QY 945 GTTTTCGCTATAGAGACTCGGTTATCCGAAGCCCGCATCTACTCTGATTTCTAGAAC 1004
DB 843 TTTTCGAGATTTGAAATTCAGCTATTAGAGTCCCACTTTAGGATTTCTTAATAA 902
QY 1005 ACTTACAAATTTTACGACTTTCATCAGCATGGAGTCTACTAGGCATATGACTTACTGGCG 1064
DB 903 TATAATTTATTGACACTGATTTA-----ATTAGAGCGGTTCACATTGGGC 947
QY 1065 GGGGCACACAAATTCATCTCGGCCAATAGAGGGCGGATTAATACTCAAGCGATGGGTC 1124
DB 948 GGGGCATCGGTGTAATCTCTCATTTTACCGGTAGTTGCAAGTGTATAAGCTCCCTCAATA 1007
QY 1125 TACCAATCTCTTATTAATCTCTGTAAGATTAATCTCTCTCGAGAGCTATATTGAC 1184
DB 1008 CGGATTAACCTCAACCGCAGAACCGAGTTCGAATATTGCTCTAGCACTTTTCCAGGTCT 1067
QY 1185 TGAATCATATGACGAGTGTCTTATGGGAAATTTACCTTGAACCTATTTCATGTGTCCC 1244
DB 1068 TAATCTATTTATAGAACACTATCAGACCCTTCTTCCGAAGATCCGATATATTATGCC 1127
QY 1245 TACTGTTAGATTTTAATTTTATAGAACCTCTCAGAAATACCTTTTGAAGAGGTACTGCTAACTA 1304
DB 1128 AACATTAGGAATAAATGTAGTGCAGGGGGTAGGATTCATTCAACCAAAATAATGGTGAAGT 1187
QY 1305 TAGTCAACCTTATGAGTCACTGGCTTCAATTAAGAGATTCAGAACTGAATTAACACC 1364
DB 1188 TCTATATAGAGGAGAGGAACAGTAGATTCTTTGATGAGTTGCCAATTGA-----CGG 1241
QY 1365 AGAAACAACAGAACGACAAATTAATGAATCATATAGTATAGTGTATCTCACATAGGGCT 1424
DB 1242 TGAGAAATTCATTTAGTGGATATAGTCAATAGATTAAGTTCAGCTTACATTAACAGGTCTTT 1301
QY 1425 CATTTCAATCTAGGGTGCATGACAGTATATTTCTTGGACGCCACCGTAGTGAGATCG 1484
DB 1302 ATATAATACATAATACTAGTGTGCCAACATTTGTTTGGACACATCACAGTGTACTGA 1361
QY 1485 TACAAATACCATTTAGTTCAGATGACATAACAAATACCATTTGTAATAATCATTTCAACCT 1544
DB 1362 TCGAAATATAATCTATCCGATGTAATTAACAAATACCATTTGTAATAATCATTTCTCCCT 1421
QY 1545 TAATTGAGTACCTCTGTAGTGCAGGCCAGGATTTACAGGAGGGGATATAATCCGAAC 1604
DB 1422 TACTTCAGGTACCTCTGTAGTCAGAGGCCAGGATTTACAGGAGGGGATATCATCCGAAC 1481
QY 1605 TAACGTTTAATGGTATGATCTAAGTATGGTCTTAATTTTAAATAATACATCATTTACAGCG 1664
DB 1482 TAACGTTTAATGGTAAATGCTACTAGTATGAGTCTTTAAATTTTAGTAAATACATCATTTACAGCG 1541
QY 1665 GTATCGCGTGAGAGTTCGTTATGCTGCTTCTCAAAACAAATGGTCTGAGGTAATCTGTGCG 1724
DB 1542 GTATCGCGTGAGAGTTCGTTATGCTGCTTCTCAAAACAAATGGTCTCATGAGAGTAAATGTTGG 1601
QY 1725 AGGAGTACTACTTTTTCATCAAGGATTTCCCTAGTACTATGATGTCGAATGAGTCTTTGAC 1784
DB 1602 AGGAGTACTACTTTTTCATCAAGGATTTCCCTAGTACTATGATGTCGAATGAGTCTTTGAC 1661
QY 1785 ATCTCAATCAATTTAGATTTGAGAAATTTCTGTAGGTAATGATGTCATCTGGCAGTCAAAC 1844
DB 1662 ATCTCAATCAATTTAGATTTGAGAAATTTCTGTAGGTAATGATGTCATCTGGCAGTCAAAC 1721
QY 1845 TGCTGGAATAAGTATAAGTAAATATGAGGAPAGCAAAACGTTTTCATCTTTGATAAATTTGA 1904
DB 1722 TGCTGGAATAAGTATAAGTAAATATCAGGTAGACAAACGTTTTCATCTTAGATAGAAATTTGA 1781
QY 1905 ATTCATTTCCAAATTAAGTAAATATGAGGAGATAGATTTAGAAAGGGCGGAGAGGC 1964
DB 1782 ATTTATCCCAAGTTGATGCAACATTTGAGGAGAGATATGATTTAGAAAGAGACAAAAAGGC 1841
QY 1965 GGTGAATGCTGTGTTTACTTAATACGAATCCAAAGAGATTTGAAACAGATGTGACAGATTA 2024
DB 1842 GGTGAATGCTGTGTTTACTTCTCCATCAATCAATCGAGTTTAAACAGATGTGACGATTA 1901
QY 2025 TCATATTGATCAAGTATCCAAATTTAGTGGCGGTGTTTATCGGATGAATTTCTGCTTAGATGA 2084

DB 1902 TCATATTGATCAAGTATCCAAATTTAGTAGATTGTTTATCCGATGAATTTGCTCTGGATGA 1961
QY 2085 AAAGAGAGAAATTTACTTTGAGAAAGTGAATATGCGAAACGACTCAGTGATGAAGAAACTTT 2144
DB 1962 AAAGCGAGAAATTTGTCGAGAAAGTCAAAATGCGAAGCGACTCAGTGATGAGCGGAATTT 2021
QY 2145 ACTCCAGATCCAAACTTTCACATCCATCAATTAAGCAACCGAGCTTCATATCTACTAATGA 2204
DB 2022 ACTTCAAGATCCAAACTTTCAGAGGATCAATAGGCAACCGAG----- 2064
QY 2205 GCAATCGAATTTCACTATCTATCCATGAACAACTCTGAACATGGATGGTGGGGAAGTGAGAA 2264
DB 2065 -----CGTGGCTGGAGAGGAAGTACCGA 2087
QY 2265 CATTTACAAATCCAGGAAGAAATGACGTATTTAAAGAGAAATTTACGTCACTACCTCCGGGAC 2324
DB 2088 TATTACCATCCAGGAGGAGATGACGTATTTCAAAGAGAAATTAAGTCACTACCTACCGGTAC 2147
QY 2325 TTTTATGAGTGTATCCGACGTATTTTATATCAAAAAATAGGAGAGTCCGNAATTTAAAGC 2384
DB 2148 CTTTGTAGTGTCTATCCAACTGTTTGTATCAAAAAATAGTGTAGTCCGAAATTTAAAGC 2207
QY 2385 TTATACTCGCTACCAATTTAAGAGGGTATATTGAAGATAGTCAAGATTTAGAGATATATTT 2444
DB 2208 CTATTAACCGTTTCAATTTAAGAGGGTATATCGAAGATAGTCAAGACTTAGAATCTATTT 2267
QY 2445 GATTCGTTTAAATGCGAAACATGAAACATTTGGATGTTTCCAGGTACCGAGTCCGTTATGCC 2504
DB 2268 AATTGCTCAATGCAAAACACGAAACAGATAATGTACCAGGTACCGGTCTCTTTATGCC 2327
QY 2505 GCTTTGAGTTGAAAGCCCAATCCGAGAGTCCGAGAAACCGAATCGATCGCACACATTT 2564
DB 2328 GCTTTGAGTCCAAATTTGAGAGTGTGAGAAACCGAATCGGTGTGTGCCACACCT 2387
QY 2565 TGAATGAAATCTGATCTAGATTTCTCTCGAGAGATGGAAGAAATGTCGCAATCTATTC 2624
DB 2388 TGAATGAAATCTGATTTAGATTTCTCTCGAGAGCGGGAAGAAATGTCGCAATCTATTC 2447
QY 2625 CCATCATTTCTCTTTGGATATTGATATTGGATGACAGACTTTCATGAGAAATCTAGGCGT 2684
DB 2448 CCATCATTTCTCTTTGGACATTTGATTTGGATGACAGACTTTCGCAAGAGGATCTAGGCGT 2507
QY 2685 GTGGGTGATTTCAAGATTAAGACGCGAGGAGGTTCATGCAAGCTAGGGAATCTGGAATTT 2744
DB 2508 GTGGGTGATTTCAAGATTAAGACGCGAGGAGGTTCATGCAAGCTAGGGAATCTGGAATTT 2567
QY 2745 TATTGAGAGAAACCATTTATTAGGAGAGCACTGCTCTGTTGAAGAGAGCAGAGAGAAAA 2804
DB 2568 TATCGAAGAGAAACCATTTATTGGAGAGCACCTGCTCTGTTGAAGAGAGCGGGAAGAAAA 2627
QY 2805 ATGGAGAGACAAACGCTGAAACAACTACAATTTGGAACAAACAGATATATACAGAGGCAAA 2864
DB 2628 ATGGAGAGACAAACGCGAAGAACTACAATTTGGAACAAACAGATATATACAGAGGCAAA 2687
QY 2865 AGAAGCTGTGGATGCTTTTATTTGTAGATTTCTCAATATAATAGATTAAGCGGATACAAA 2924
DB 2688 AGAAGCTGTGGATGCTTTTATTTGTAGATTTCTCAATATAATAGATTAAGCGGATACAAA 2747
QY 2925 CATTTGCAATTCATGCGGCAGATAAATGTTGTTTCATGCAATTCGAGAGGCTTATCTGTC 2984
DB 2748 CATTTGATGATTCATGCGGCAGATAGACTGTTGTTTCATGCAATTCGAGAGGCTTATCTTCC 2807
QY 2985 AGAATTTATCTGTTATCCCGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAAAGTCCGAT 3044
DB 2808 AGAATTTATCTGTTATCCCGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAAAGTCCGAT 2867
QY 3045 TATCACTGCAATCTCCCTATACGATCGGAAATGTCGTTTAAATAATGCTGATTTTAAATAA 3104
DB 2868 TTCTACTGCAATTTCCCTATATGATGCGAGAAATGTCATTTAAATAATGCGGATTTCAATAA 2927
QY 3105 TGAATTTAGCATGCTGGATGTAAAGGGCATGTAGAT---GTACAAACAGAGCCATCACCG 3161

```

Db 2928 TGGCTTATCATGCTGGAAACGTAAGAGGCATGTAGATGTAGTACAAACAAACCCCG 2987
Qy 3162 TTCTGTCTCTCTGTATCCAGAAATGGAGACAGAGTGTCAACAGCAGTTCCGCTCTGTCC 3221
Db 2988 TTCCGTCTCTCTGTATCCGGAATGGAGACAGAGTGTCAACAAATTTCTGTCTGTCC 3047
Qy 3222 GGGCGTGGCTATATCTCTCTGTACAGCGTACAAAGAGGGATATGGAGAGGTTGTGT 3281
Db 3048 GGGCGTGGCTATATCTCTCTGTACAGCGTACAAAGAGGGATATGGAGAGGTTGTGT 3107
Qy 3282 AACATCCATGAAATCGAGAACATACAGACGAACTAAATTTAAACCTGTGAAGAAGA 3341
Db 3108 AACATCCATGAGATCGAGAACATACAGACGAACTAAATTTAAACCTGTGAAGAAGA 3167
Qy 3342 GGAAGTGTATCCACGAGATACAGAACTGTGAATGATTATCTCCACACCAAGGTACAGC 3401
Db 3168 GGAAGTGTATCCACGAGATACAGAACTGTGAATGATTATCTCCACACCAAGGTACAGC 3227
Qy 3402 AG-----TATGTAATTCCTGTAATGCTGGATATGAGGATGCATATGAAGTTGA 3449
Db 3228 AGGATCCACAGATTCATGTAATTCCTGTAATGCTGGATATGAGGATGCATATGAAGTGA 3287
Qy 3450 TACTACAGCATCTGTAATTAACACCGACTTATGAAGAGAAACCTATACAGATGTACG 3509
Db 3288 TACTACAGCATCTGTAATTAACACCGACTTATGAAGAGAAACCTATACAGATGTACG 3347
Qy 3510 AAGAGATAATCATGTGAATGATACAGAGGATGTGAATTTATCCACACTACAGCTGG 3569
Db 3348 AGGAGATATCATGTAATTAACACCGACTTATGAAGAGAAACCTATACAGATGTACG 3407
Qy 3570 TTATATGACAAAGAAATAGAAATCTCCAGAAACCGATTAAGATGTAGATGTGAGTTGG 3629
Db 3408 TTATGTGACAAAGAAATAGAGTACTTCCAGAAACCGATTAAGATGTAGATGTGAGTTGG 3467
Qy 3630 AGAACCGAGGAGGAGTTATTTAGACAGCGTGAATTTACTCTTATGAGGAA 3684
Db 3468 AGAACCGAGGAGGAGTTATTTAGACAGCGTGAATTTACTCTTATGAGGAA 3522

```

```

RESULT 14
PCT-US95-10310-1
; Sequence 1, Application PC/TUS9510310
; GENERAL INFORMATION:
; APPLICANT: MYCOGEN CORPORATION
; APPLICANT: STREET ADDRESS: 5501 Oberlin Drive
; APPLICANT: CITY: San Diego
; APPLICANT: STATE/PROVINCE: California
; APPLICANT: COUNTRY: US
; APPLICANT: POSTAL CODE/ZIP: 92121
; APPLICANT: PHONE NUMBER: (619) 453-8030
; APPLICANT: FAX NUMBER: (619) 453-6991
; TITLE OF INVENTION: Protein Toxins Active Against Lepidopteran Pests
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10310
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/597,607
; FILING DATE: 15-OCT-90
; CLASSIFICATION:

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA50.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (904)375-8100
; TELEFAX: (904)372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3522 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bacillus thuringiensis
; STRAIN: aizawai
; INDIVIDUAL ISOLATE: PS81A2
; IMMEDIATE SOURCE:
; LIBRARY: Lambdaagem - 11 (tm) Library of August Sick
; CLONE: 81A2
; PCT-US95-10310-1

Query Match 48.6%; Score 1793.4; DB 5; Length 3522;
Best Local Similarity 71.9%; Pred. No. 0;
Matches 2497; Conservative 0; Mismatches 861; Indels 117; Gaps 7;

Qy 228 ACAAATAGCTAGTATTTTATAGTCTTTCTGTGTGGTAATATGCGCCCGCGCAGAGATCA 287
Db 147 ACTAGGGGATTTTATATCTTGGCTTGTGTGATTAATATGGGGGCTATAGGTCTCTTCA 206
Qy 288 GTGGGAATTTCTAGAACATGTCGAACAACTTATAAATCAACAATAACAGAAATGC 347
Db 207 ATGGATATATTTTATAGCAAAATGAGCTATTGATCGGCCAAGAAATAGAGAAATTCG 266
Qy 348 TAGGAATACCGGCACTTGTCTGATTACAAAGTTTAGGAGATTCTTTTAGAGCCTATCAAC 407
Db 267 TAGGAATACGGCAATTTCTAGATTACAAAGGCTAAAGCAATCTTTACCGAATTTACACA 326
Qy 408 GTCATTTGAAGATGGCTAGAAAACCGTATGATGATGACAGAGAGAGTCTTCTTATAC 467
Db 327 TGCTTTTAAAACTGGGAGTAGATCCTACTAATCCAGCATTAAGAGAGAGATGCGTAT 386
Qy 468 CCAATATATAGCTTAGAACTTCTGATTTCTTAATGCGATCGCGCTTTTCGCAATTAGAAA 527
Db 387 TCAATTTATGACATGAACAGTCTCTTCAACAGCTATTTCTTTTTCAGTTCAAGG 446
Qy 528 CCAAGAAATTCATTTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTTATT 587
Db 447 TTATGAATTTCTCTTTTATCAGTATATGTTCAAGCTGCAAAATTTACATTTATCGGTTT 506
Qy 588 GAGAGATGCTCTCTTTTGGTATGAAATTTGGGCTTACATCGCAGGAAATTTCAAGTTA 647
Db 507 GAGAGATGTTTTCAGTGTTTGGCAACCGTTGGGATTTTGTATGTAGCAACAATATAGTC 566
Qy 648 TTATGAGCGCAAGTGGAAACAAACGAGAGATTATTCGACTATTTCGCTAGCAATGTTATA 707
Db 567 TTATATGATTTTAACTAGGCTTATTGGCGAATATCTGATTTGCTGCTAGTTGGTATAA 625
Qy 708 TACAGGTCTAAATAGCTTGAGAGGGAACAAATGCGC---CAAGTTGGGTGCTTTATATCA 764
Db 627 TACGGGGTTAAATCGTTTACACGTAATGAAGGGGTACGAGGATGGCAAGATTTAATAG 686
Qy 765 ATTCCGTAGAGATCTAACGTTAGGGTATTTAGATCTAGTGGCACTATTTCCCAAGCTATGA 824
Db 687 GTTTAGAAGAGAGGTAAACAATATCAGTATTTAGATATTTATTTCTTTTTCAAAATTTACGA 746
Qy 825 CACTCGCACCTTATCCCAATTAATACGAGTGTCTAGTTAAACAAAGGAACTTTTATACAGACGC 884
Db 747 TTCTAGATTATATCCCAATTTCCGCAATCTATCAATTAACGGGGAGTATATACAGATCC 806

```

QY 885 AATTGGACCAACAGGGTAATAATGCGCAAGTATGAAATGGGTATTAATAATGCACTTC 944
Db 807 -----GGTAATTAATAATAACTGATTATAGAGTTACCCCAAG 842
QY 945 GTTTTCGGCTATAGAGACTGGGTTATCCGAAGCCGCGATCTACTTGATTTTCTAGAAC 1004
Db 843 TTTCCGAGAGTATGAAATTCAGCTATTAGAGTCCCGCATCTTATGGATTTCTTAATAA 902
QY 1005 ACTTACAAATTTTACACTTCAATCAGATGAGGTGCTACTAGGCATATGACTTACTGGC 1064
Db 903 TATAAATTATTGACACTGATTTA-----ATTAGAGCGGTTCACTATTGGCC 947
QY 1065 GGGGCACAAATTCAACTTCGGCCCAATAGGAGCGGATTAATACTCTCAAGCGATGGTC 1124
Db 948 GGGGCATCGTGAATCTTCTTATTTTACCGGTAGTTGCGAAGTGAATAAGCTCCCTCAATA 1007
QY 1125 TACCAATACTTCTAATTAATCTGTAAGATTAATCACTTCTCTCGAGACGTATATTGGAC 1184
Db 1008 CGGATAACTGCAACGACGAGTCGAACTATTGCTCTTAGCACATTTTCCAGTCT 1067
QY 1185 TGAATCATATGACGAGGTGCTTCTATGGGGAATTTACCTTGAACCTATTCAATGGTGTCC 1244
Db 1068 TAATCTATTTTATAGAACACATCATCAGACCCCTTCTCTCGAAGATCCGATATAATTATGCC 1127
QY 1245 TACTGTTAGATTAATTTTGAACCCCTCAGAACTACTTTTGAAGAGGTACTGCTAACTA 1304
Db 1128 AACATTAGGAATNAATAGTGCAGGGGGTAGGATTCATTCAACCAATATGTTGAAGT 1187
QY 1305 TAGTCAACCCCTATGAGTCACCTGGCTTCAATTAAGAGATTCAGAACTGAATTTACCACC 1364
Db 1188 TCTATATAGAAGGAGGACAGTAGATTCTCTTGATGAGTTGCCAATTGA-----CGG 1241
QY 1365 AGAAACACAGAACGCAAAATTTAGTAACTATATAGTCAATAGGTTATCTCAATAGGGCT 1424
Db 1242 TGAAATTCATTAGTTGGATAGTCAATAGTAAAGTCAAGTCACTTAAACAGGTGCTT 1301
QY 1425 CATTTCAATCTAGGTGATGATACCGATATTTCTTGGACGACCGTAGTGAGATCG 1484
Db 1302 ATATAACTAATAATACTAGTCTGCCAACATTTGTTTGGACACATCACAGTGTACTGA 1361
QY 1485 TACAAATACCAATAGTTAGATACATAACAAATACCAATGTTGTAATCAATCAACCT 1544
Db 1362 TCGAAATATAATCTATCCGATGTAAATACAAATACCAATGTTGTAATCAATCTCCCT 1421
QY 1545 TAATTCAGGTACCTCTGTAGTCACTGCGCCAGGATTTTACAGAGGGGATATAATCCGAAC 1604
Db 1422 TACTTCAGGTACCTCTGTAGTCAGAGCCAGGATTTTACAGAGGGGATATCATCCGAC 1481
QY 1605 TAACGTTAATGGTATGTAATAGTATGGGTCTTAAATTTTAAATAATACATCATACAGC 1664
Db 1482 TAACGTTAATGGTAAATGTAATAGTATGAGTCTTAAATTTTAAATAATACATCATACAGC 1541
QY 1665 GTATCGGTGAGAGTTCGTTATGCTGCTCTCAACCAATGGTCTGAGGGTAACTGTCGG 1724
Db 1542 GTATCGGTGAGAGTTCGTTATGCTGCTCTCAACCAATGGTCTGAGAGTAAATGTTGG 1601
QY 1725 AGGAGTACTACTTTTGAATCAAGGATTCCTTAGTACTATAGTGCATAATGAGTCTTTGAC 1784
Db 1602 AGGAGTACTACTTTTGAATCAAGGATTCCTTAGTACTATAGTGCATAATGAGTCTTTGAC 1661
QY 1785 ATCTCAATCATTTAGATTTGAGAAATTTCTGTAGGTATAGTGCATCTGGCAGTCAAAC 1844
Db 1662 ATCTCAATCATTTAGATTTGAGAAATTTCTGTAGGCAATTAGTACATCTGGCAGTCAAAC 1721
QY 1845 TGCTGGAATAGTATAGTATAGTATAGGTAGACAAACGTTTCACTTTGATAAATTTGA 1904
Db 1722 TGCTGGAATAGTATAGTATAGTATAGGTAGACAAACGTTTCACTTTAGATAAATTTGA 1781
QY 1905 ATTCAATTCCAATTTACTGCAACTTTGAGAGCAATATGATTTAGAAAGGGCGCAAGGCG 1964
Db 1782 ATTTATCCAGTTGATGCAACATTTGAGAGCAATATGATTTAGAAAGGACCAAAAGGC 1841
QY 1965 GGTGAATGCTCTGTTTACTAATACGAATCCAAAGAAGATTGAAACAGATGTGACAGATTA 2024

Db 1842 GGTGAATTCGCTGTTTACTTCTTCCAATCAAATCGAGTTAAAAACAGATGTGACGGATTA 1901
QY 2025 TCATATTGATCAAGTATCCAAATTTAGTGGCGTGTATTTCGATGAAATTTCTCTTAGATGA 2084
Db 1902 TCATATTGATCAAGTATCCAAATTTAGTAGATTTGTTTATCCGATGAAATTTCTCTGATGA 1961
QY 2085 AAAGAGAGATTTACTTGGAGAAAGTGAATATGCGAAACGACTCAGTGTATGAAGAACTT 2144
Db 1962 AAAGCGAGAAATTTGTCGAGAAAGTCAAAACATGCGAAGCGACTCAGTGTATGAGCGGAAATTT 2021
QY 2145 ACTCCAGATCCAAACTTTCATCCATCAATAAAGCAACAGACTTCATATCTACTAATGA 2204
Db 2022 ACTTCAAGATCCAAACTTTCAGAGGATCAATAGGCAACAGAC----- 2064
QY 2205 GCAATCGAATTTTCATCTATCCATGAACAATCTGAAACATGATGGTGGGGAAGTGAGAA 2264
Db 2065 -----CGTGGCTGGAGAGGAGTACGGA 2087
QY 2265 CATTAATTCAGAGAGAAATGACGTATTTAAAGAGAAATTTACGTACACTACCGGGAC 2324
Db 2088 TATTACCATCCAGAGGAGATGACGTATTTCAAGAGAAATTTACGTACACTACCGAGTAC 2147
QY 2325 TTTTAAATGAGTGTATCCGACGTATTTATATCAAAAATAGGAGAGTCGGAATTTAAAGC 2384
Db 2148 CTTTGTAGTGTCTATCCACGTATTTGTATCAAAAATAGTGTGAGTGTGGAATTTAAAGC 2207
QY 2385 TTATCTCGCTTACCAATTAAGAGGGTATATTGAAGATAGTCAAGATTTAGAGATATATTT 2444
Db 2208 CTATACCGTTTCAATTAAGAGGGTATATCGAGATAGTCAAGACTTAGAATCTATTT 2267
QY 2445 GATTGCTTAAATGCGAAACATGAAACATTTGGATGTTCCAGGTACCGAGTCCGTATGGCC 2504
Db 2268 AATTGCTCTAATGCAAAACACGAAACAGTAAATGTACAGGTACGGGTCTCTTATGGCC 2327
QY 2505 GCTTTCAATTTGAAGCCCAATCGGAGAGTGGGAGACCGAATCGATGCGCACACATTT 2564
Db 2328 GCTTTCAATCGAAAGTCCAAATTTGGAAGGTGTGGAGAACCGAATCGGTGTGTGCGCACACCT 2387
QY 2565 TGAATGAAATCTCTGATCTAGATTTGCTCTGAGAGATGGAGAAATTTGTGGCATCATTC 2624
Db 2388 TGAATGAAATCTCTGATTTAGATTTGCTCTGAGAGACGGGGAATTTGTGCACATCATTC 2447
QY 2625 CCATCATTTCTTTTGGATATGATTTGATGACAGACTTTGATGAGAAATCTAGGCGT 2684
Db 2448 CCATCATTTCTCTTGGACATTTGATTTGGATGACAGACTTTGCAAGAGGATCTAGGCGT 2507
QY 2685 GTGGTGTATTTCAAGATTAAGACCGAGGAGGTTCATGCAAGACTAGGGATCTGGAAAT 2744
Db 2508 GTGGTGTATTTCAAGATTAAGACCGAGGAGGTTCATGCAAGATTTAGGAAATCTGGAAAT 2567
QY 2745 TATTGAAGAGAAACCATTTATTAGGAGAGCACTCTCTGTTGAAAGAGAGAGAGAGAGAGAG 2804
Db 2568 TATCGAAGAGAAACCATTTATTGGAGAGCACTCTCTCTGTTGAGAGAGAGAGAGAGAGAG 2627
QY 2805 ATGAGAGAGACAAACGTTGAAACACTTCAATTTGGAACAAACAGAGTATATACAGAGGCAA 2864
Db 2628 ATGAGAGAGACAAACGTTGAAACACTTCAATTTGGAACAAACAGAGTATATACAGAGGCAA 2687
QY 2865 AGAAGCTGTGATGCTTTTATTTGTAGATTTCTCAATATAATAGATTTACAGCGGATACAA 2924
Db 2688 AGAAGCTGTGATGCTTTTATTTGTAGATTTCTCAATATAATAGATTTACAGCGGATACAA 2747
QY 2925 CATTTGGCATGATTCATCGGCGAGATAAATTTGTTTCATCGAATTCGAGAGGCTTATCTGTC 2984
Db 2748 CATTTGGTATGATTCATCGGCGAGATAGACTTTGTTTCATCGATCCAGAGGCTTATCTCC 2807
QY 2985 AGAATTAATCTGTTATCCCGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAGGTGCGAT 3044
Db 2808 AGAATTAATCTGTTATCCAGGAAATTAATGTTGTTGATTTTGAAGAAATTTAGAAACCGTAT 2867
QY 3045 TATCACTGCAATCTCCCTATACGATGCGAGAAATGTGCTTTAAAAATCGTGTATTTTATATA 3104

Db 2868 TTCTACTGCAATATCCCTATATGATGCGAGAAATGTCTAATTAATAATGGCGATTTCAATAA 2927
 Qy 3105 TGGATTAGCATGCTGGAATGTAAAGGCGCATGTAGAT---GTACAACAGAGCCATCACCG 3161
 Db 2928 TGGCTTATCATGCTGGAACGTGAAAGGCGATGTAGATGTAGTAAGAAACACACACCG 2987
 Qy 3162 TTCTGTCTCTTGTATCCAGAAATGGGAAGCAGAGAGTGTCAACAAGCATGTGCGTCTGTCC 3221
 Db 2988 TTCGTCTCTTGTGTCGCGGAATGGGAAGCAGAGAGTGTCAACAACAATTCGTGTCTGTCC 3047
 Qy 3222 GGGGCGTGGCTATATCTCCGCTGTACAGCGTACAAAGAGGATATGGAGGCTGTCTGT 3281
 Db 3048 GGGGCGTGGCTATATCTCCGCTGTACAGCGTACAAAGAGGATATGGAGAGGTTCGGT 3107
 Qy 3282 AACGATCCATGAAATCGAGAAACAATACAGACGAACCTAAATAATTTAAATACTGTGAAGAAGA 3341
 Db 3108 AACCATCATGAGATCGAGAACAAATACAGACGAACCTAAATAATTTAAATACTGTGAAGAAGA 3167
 Qy 3342 GGAAGTGTATCCAAACGGATACAGAAACGTGTAAATGATTTATCTGCAACCAAGGTACAGC 3401
 Db 3168 GGAAGTGTATCCAAACGGATACAGAAACGTGTAAATGATTTATCTGCAACCAAGGTACAGC 3227
 Qy 3402 AG-----TATGTAATTTCCGTAATGCTGGATATCAGGATGCAATATGAAGTTGA 3449
 Db 3228 AGGATCCACAGATTCATGTAATTCCTGTAATATCAGATATGAGGATGCAATATGAATGAA 3287
 Qy 3450 TACTACAGCATCTGTAATTAACAACCGACTTATGAAGAGAAACGCTATACAGATGTACG 3509
 Db 3288 TACTACAGCATCTGTAATTAACAACCGACTTATGAAGAGAAAGGTATACAGATGTACA 3347
 Qy 3510 AAGAGATAATCATTGTGAATATGACAGAGGTATGTGAATATCCACCTACCAAGCTGG 3569
 Db 3348 AGGAGATAATCATTGTGAATATGACAGAGGTATGTGAATATCCAGCAGTACCAGCTGG 3407
 Qy 3570 TTATATGACAAAGAAATAGNAATCTCCAGAAACCGATAGATGATGATGATGATG 3629
 Db 3408 TTATGTGACAAAGAAATAGAGTACTTCCCAAGAAACCGATAGATGATGATGATGATG 3467
 Qy 3630 AGAAACGGAAGGAAGTTTATTTGAGACAGCGTGAATATCTCTTATGAGGAA 3684
 Db 3468 AGAAACGGAAGGAAGTTTATTTGAGACAAATGTCGAATTACTCTTATGAGGAA 3522

RESULT 15
 5164180-3
 ; Patent No. 5164180
 ; APPLICANT: Payne, Jewel; Sick, August J.
 ; TITLE OF INVENTION: BACILLUS THURINGIENSIS ISOLATES ACTIVE
 ; AGAINST LEPIDOPTERAN PESTS
 ; NUMBER OF SEQUENCES: 6
 ; CURRENT APPLICATION DATA:
 ; FILING DATE: 14-DEC-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 353,860
 ; FILING DATE: 18-MAY-1989
 ; SEQ ID NO: 3
 ; LENGTH: 3522
 5164180-3

Query Match 48.6%; Score 1791.8; DB 6; Length 3522;
 Best Local Similarity 71.8%; Pred. No. 0;
 Matches 2496; Conservative 0; Mismatches 862; Indels 117; Gaps 7;
 Qy 228 ACAATAGCTAGTTTATATGTTTCTGTTGGTGAATATGCGCCCGCGCAGAGATCA 287
 Db 147 ACTAGGGGATTTATATCTGCTTGTGATGTAATATGGGGGCTATAGGTCCTTCACA 206
 Qy 288 GTGGGAAATTTTCTAGAAACATGTCGAACAACTTTATAAATCAACAAATAACAAATATGC 347
 Db 207 ATGGGATATATTTTAGGCAAAATGAGCTATGATCGGCCAAAGATAGAGGAATTCG 266
 Qy 348 TAGGAATACGGCACTTCTGCTGATATACAAAGTTTAGGAGATTCCTTTAGAGCCTATCAACA 407

Db 267 TAGGAATCAGGCAATTTCTAGATTTACAAGGCTAAAGCAATCTTTACGAAATTTACACAAA 326
 Qy 408 GTCACTTGAAGATGGCTAGAAAACCGTATGATGCAAGAACGAGAGTGTCTTTATATAC 467
 Db 327 TGTCTTTTAAAAACCTGGGAAGTAGATCTCTACTAATCCAGCATTAAGAGAAGATGGCTAT 386
 Qy 468 CCAATATATAGCCCTTAGAACTTGATTTCTTAATGCGATCCGCTTTTCGCAATATAGAAA 527
 Db 387 TCAATTTAATGACATGAACAGTCTCTTACAACAGCTATTTCTCTTTTTCAGTTCAAGG 446
 Qy 528 CCAAGAAAGTTCCATTTAATATGTTATATGCTCAAGCTGCAAAATTTACCTATTTATTT 587
 Db 447 TTATGAAATTTCTCTTTTATCAGTATATGTTCAAGCTGCAAAATTTTACATTTATCGGTTT 506
 Qy 588 GAGAGATGCTCTCTTTTGGTAGTGAATTTGGCTTACATCGCAGGAAATTTCAACGTTA 647
 Db 507 GAGAGATGTTTCAGTGTTTGGACAAGCTTGGGATTTGATGTAGCAACAATCAATAGTCG 566
 Qy 648 TTATGAGCGCAAGTGGAAACAAACGAGAGATTTTCGACTATTTGCTAGAAATGGTATAA 707
 Db 567 TTATATGATTTAACTAGGCTTATTTGGCAATATCTGATTTATGCTGACGTTGGTATAA 626
 Qy 708 TACAGGCTTAATAGCTTTGAGAGGCAAAATGCGG---CAAGTTGGTGGCTTTATATCA 764
 Db 627 TACGGGTTAAATCGTTTACCAGTAATGAAGGGTACGAGGATGGGCAAGATTTAATAG 686
 Qy 765 ATTCCGTAGAGATCTAACGTTAGGGGTATTTAGATCTAGTGGCACTATTTCCCAAGCTATGA 824
 Db 687 GTTTAGAAGAGAGTTAAACAATATCAGTATTTAGATATTTATTTCTTTTCCAAAATTTACGA 746
 Qy 825 CACTCGCACTTATCCCAATAAATACGAGTGTCTCAGTTTAAACAAGGGAATTTTATACAGACG 884
 Db 747 TTCTAGATTTATCCAAATTTCCGACAACTCTATCAATTTAACGCGGGAAGTATATACAGATCC 806
 Qy 885 AATTGGAGCAACAGGGGTAAATATGCGAAGTATGAATTTGGTATAAATAATGCACTTC 944
 Db 807 -----GGTAATTAATAACTGATTATAGATTATACCCCAAG 842
 Qy 945 GTTTTCGCTATAGAGATGCGGTATTCGGAACCGCATCTACTTTGATTTTCTAGAAACA 1004
 Db 843 TTTGAGAGATTTGAAATTTCACTATTTAGAGTCCCATCTTTATGGAATTTCTTAAATAA 902
 Qy 1005 ACTTACAATTTTATGACATTTCACTACGATGGAGTGTCTACTAGGCATATGACTTTACTGGG 1064
 Db 903 TATAATTTTGCACATGATTTA-----ATTAGAGGGGTTTCACTATTGGGC 947
 Qy 1065 GGGGCACACAATTTCAATCTCGGCCAATAGAGGCGGATTAATACTCAACGATGGGTC 1124
 Db 948 GGGGCATCGGTAACTTCTCATTTTACCGGTAGTTCCGCAAGTGAAGTCCCTCAATA 1007
 Qy 1125 TACCAATACCTTTATTAATCTCTGAAGATTTATCATTTCTCTCGAGACCTATATTGGAC 1184
 Db 1008 CGGGAATACGCAACCGCAGAACCGAGTGAACATTTGCTCCTAGCACTTTTCCAGGTCT 1067
 Qy 1185 TGAATCATATGCAAGGATGCTCTATGGGAAATTTTACCTTGAACCTTATCATGGTGTCC 1244
 Db 1068 TAATCTATTTTATAGAACAATATCAGACCTCTTTCCGAAGATCCGATAATATTATGCC 1127
 Qy 1245 TACTGTTAGATTTAATTTTAGGAAACCTCAGATACTTTTGAAGAGGATCTCTAATACTA 1304
 Db 1128 AACATTAGGAATAAATGTAGTGCAGGGGTAGGATTTCAATTCACCAAAATAATGGTGAAGT 1187
 Qy 1305 TAGTCAACCTTATGAGTCACTCGGCTTCAATTTAAAGATTTCAAGAACTTGAATATACCACC 1364
 Db 1188 TCTATATAGAGAGAGAGAACAGTAGATTTCTTTGATGAGTTCGCAATTTGA-----CGG 1241
 Qy 1365 AGAAACAACAGACAGCAACCAATTTATGAATCATATAGTCTATAGGTTATCTCACATAGGGCT 1424
 Db 1242 TGAGAATTCATTAGTTGGATATAGTATAGATTTAAGTCAAGTTTACATTAACCAAGGTCGT 1301
 Qy 1425 CATTTCACAATCTAGGGTGTGATGTACAGTATATTTCTTGGACGCACTGTAGTCAGATCG 1484

Search completed: April 21, 2005, 03:37:49
Job time : 635 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 20, 2005, 19:48:34 ; Search time 1933 Seconds

(without alignments)
11571.632 Million cell updates/sec

Title: US-10-614-524-1

Perfect score: 3687

Sequence: 1 ttgactcaaataggaaaa.....tactcttatggagggaatag 3687

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5622541 seqs, 3033355566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3687	100.0	3687	17 US-10-614-524-1	Sequence 1, Appli
2	3666.2	99.4	4173	17 US-10-428-961-37	Sequence 37, Appl
3	3260.2	88.4	3684	17 US-10-428-961-62	Sequence 62, Appl
4	3133.6	85.0	3687	18 US-10-809-953-9	Sequence 9, Appli
5	1581.6	42.9	3558	9 US-09-826-660-22	Sequence 22, Appl
6	1495.6	40.6	3624	10 US-09-988-462-6	Sequence 6, Appli
7	1408.4	38.2	3522	9 US-09-826-660-5	Sequence 5, Appli
8	1408.4	38.2	3522	11 US-09-837-961-7	Sequence 7, Appli
9	1408.4	38.2	3522	18 US-10-825-751-7	Sequence 25, Appl
10	1402	38.0	3534	9 US-09-873-873-25	Sequence 25, Appl
11	1402	38.0	3534	10 US-09-916-956A-25	Sequence 25, Appl

12	1402	38.0	3534	10 US-09-997-914-25	Sequence 25, Appl
13	1402	38.0	3534	16 US-10-365-645-25	Sequence 25, Appl
14	1402	38.0	3534	17 US-10-672-163-25	Sequence 25, Appl
15	1402	38.0	3534	18 US-10-739-482-25	Sequence 25, Appl
16	1402	38.0	3534	18 US-10-817-182-25	Sequence 25, Appl
17	1400.6	38.0	3531	9 US-09-873-873-9	Sequence 9, Appli
18	1400.6	38.0	3531	9 US-09-873-873-11	Sequence 11, Appl
19	1400.6	38.0	3531	9 US-09-873-873-13	Sequence 13, Appl
20	1400.6	38.0	3531	10 US-09-916-956A-9	Sequence 9, Appli
21	1400.6	38.0	3531	10 US-09-916-956A-11	Sequence 11, Appl
22	1400.6	38.0	3531	10 US-09-916-956A-13	Sequence 13, Appl
23	1400.6	38.0	3531	10 US-09-997-914-9	Sequence 9, Appli
24	1400.6	38.0	3531	10 US-09-997-914-11	Sequence 11, Appl
25	1400.6	38.0	3531	10 US-09-997-914-13	Sequence 13, Appl
26	1400.6	38.0	3531	16 US-10-365-645-9	Sequence 9, Appli
27	1400.6	38.0	3531	16 US-10-365-645-11	Sequence 11, Appl
28	1400.6	38.0	3531	16 US-10-365-645-13	Sequence 13, Appl
29	1400.6	38.0	3531	17 US-10-672-163-9	Sequence 9, Appli
30	1400.6	38.0	3531	17 US-10-672-163-11	Sequence 11, Appl
31	1400.6	38.0	3531	17 US-10-672-163-13	Sequence 13, Appl
32	1400.6	38.0	3531	18 US-10-739-482-9	Sequence 9, Appli
33	1400.6	38.0	3531	18 US-10-739-482-11	Sequence 11, Appl
34	1400.6	38.0	3531	18 US-10-739-482-13	Sequence 13, Appl
35	1400.6	38.0	3531	18 US-10-817-182-9	Sequence 9, Appli
36	1400.6	38.0	3531	18 US-10-817-182-11	Sequence 11, Appl
37	1400.6	38.0	3531	18 US-10-817-182-13	Sequence 13, Appl
38	1392.4	37.8	3534	9 US-09-873-873-27	Sequence 27, Appl
39	1392.4	37.8	3534	10 US-09-916-956A-27	Sequence 27, Appl
40	1392.4	37.8	3534	10 US-09-997-914-27	Sequence 27, Appl
41	1392.4	37.8	3534	16 US-10-365-645-27	Sequence 27, Appl
42	1392.4	37.8	3534	17 US-10-672-163-27	Sequence 27, Appl
43	1392.4	37.8	3534	18 US-10-739-482-27	Sequence 27, Appl
44	1392.4	37.8	3534	18 US-10-817-182-27	Sequence 27, Appl
45	1390	37.7	3567	10 US-09-972-175-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1

US-10-614-524-1
; Sequence 1, Application US/10614524
; Publication No. US20040016020A1
; GENERAL INFORMATION:
; APPLICANT: Arnaut, Greta
; APPLICANT: Boets, Annemie
; APPLICANT: Damme, Nicole
; APPLICANT: Mathieu, Eva
; APPLICANT: Vanneste, Stijn
; APPLICANT: Van Rie, Jeroen
; TITLE OF INVENTION: Insecticidal proteins from *Bacillus thuringiensis*.
; FILE REFERENCE: NEWBTSUS2
; CURRENT APPLICATION NUMBER: US/10/614,524
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: US/09/739,243
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/173387
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 1
; LENGTH: 3687
; TYPE: DNA
; ORGANISM: *Bacillus thuringiensis*
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3687)
US-10-614-524-1

Query Match 100.0%; Score 3687; DB 17; Length 3687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGACTTCAATAGAGAAATGAGATGAAATTAATAATGCTTTATTCGATTTCAGCTGTA 60
DB 1 TTGACTTCAATAGAGAAATGAGATGAAATTAATAATGCTTTATTCGATTTCAGCTGTA 60
QY 61 TCGAATCATTTCCACACAAATGGATCTATACCCAGATGCTCGTATTTGAGGATTTCTTTGT 120
DB 61 TCGAATCATTTCCACACAAATGGATCTATACCCAGATGCTCGTATTTGAGGATTTCTTTGT 120
QY 121 ATAGCCGAGGGGATATATCAATCCACTTTGTTAGCGCATCAACAGTCCCAACGGGTATT 180
DB 121 ATAGCCGAGGGGATATATCAATCCACTTTGTTAGCGCATCAACAGTCCCAACGGGTATT 180
QY 181 AACATAGCTGCTAGAACTAGGTGATTTAGCGGTACCGTTTGTGACACAAATAGCTAGT 240
DB 181 AACATAGCTGCTAGAACTAGGTGATTTAGCGGTACCGTTTGTGACACAAATAGCTAGT 240
QY 241 TTTTATAGTTTCTTTGTTGTAATTAATGCCCCCGCGCAGAGATCAGTGGGAAATTTTC 300
DB 241 TTTTATAGTTTCTTTGTTGTAATTAATGCCCCCGCGCAGAGATCAGTGGGAAATTTTC 300
QY 301 CTAGAACATGTCGAACTTATTAATCAACAAATACAGAAATGCTAGGAATACGGCA 360
DB 301 CTAGAACATGTCGAACTTATTAATCAACAAATACAGAAATGCTAGGAATACGGCA 360
QY 361 CTTGCTCGATTACAAGTTTATAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTTGAAGAT 420
DB 361 CTTGCTCGATTACAAGTTTATAGGAGATTCCTTTAGAGCCTATCAACAGTCACTTTGAAGAT 420
QY 421 TGGCTAGAAACCGTGATGATGCAAGAAACGAGAGTGTCTTTATACCCCAATATATAGCC 480
DB 421 TGGCTAGAAACCGTGATGATGCAAGAAACGAGAGTGTCTTTATACCCCAATATATAGCC 480
QY 481 TTAGAACTTGATTTCTTTAATGCGATGCGCTTTTCGCAATTAGAACCAAGATTTCCA 540
DB 481 TTAGAACTTGATTTCTTTAATGCGATGCGCTTTTCGCAATTAGAACCAAGATTTCCA 540
QY 541 TTATTAATGCTATGCTCAAGCTGCAAAATTTACACCTATTATTATTTAGAGAGATGCTCT 600
DB 541 TTATTAATGCTATGCTCAAGCTGCAAAATTTACACCTATTATTATTTAGAGAGATGCTCT 600
QY 601 CTTTTTGGTATGAAATTTGGGCTTACATCGCAGGAAATTCAGCTTATTAATGAGCGCCAA 660
DB 601 CTTTTTGGTATGAAATTTGGGCTTACATCGCAGGAAATTCAGCTTATTAATGAGCGCCAA 660
QY 661 GTGCAACAAACGAGAGATTTTCGACTATTTGGGTAGAAATGGTATATACAGGTCTAAAT 720
DB 661 GTGCAACAAACGAGAGATTTTCGACTATTTGGGTAGAAATGGTATATACAGGTCTAAAT 720
QY 721 AGCTTGAGAGGGCAAAATGCGCAAGTTGGGTGCTTATTAATCAATTCGCTAGAGATCTA 780
DB 721 AGCTTGAGAGGGCAAAATGCGCAAGTTGGGTGCTTATTAATCAATTCGCTAGAGATCTA 780
QY 781 ACGTTAGGGGTATAGATCTAGTGGCACTATTTCCCAAGCTATGACACTCGCACTTATCCA 840
DB 781 ACGTTAGGGGTATAGATCTAGTGGCACTATTTCCCAAGCTATGACACTCGCACTTATCCA 840
QY 841 ATAAATACGAGTCTCAGTTTAAACAGGGAAGTTTATACAGACCAATTTGAGCAACAGGG 900
DB 841 ATAAATACGAGTCTCAGTTTAAACAGGGAAGTTTATACAGACCAATTTGAGCAACAGGG 900
QY 901 GTAAATATGCAAGTATGAAATTTGGTATTAATAATCACTTCGTTTTCGGTATATAGAG 960
DB 901 GTAAATATGCAAGTATGAAATTTGGTATTAATAATCACTTCGTTTTCGGTATATAGAG 960
QY 961 ACTGCGGTATCCGAAGCCGCACTACTTGAATTTCTAGAACAACTTACAAATTTTAGC 1020
DB 961 ACTGCGGTATCCGAAGCCGCACTACTTGAATTTCTAGAACAACTTACAAATTTTAGC 1020
QY 1021 ACTTCATCAGATGGAGTGTCTAGGCAATATGACTTACTGCGGGGGGACAAATTTCAA 1080
DB 1021 ACTTCATCAGATGGAGTGTCTAGGCAATATGACTTACTGCGGGGGGACAAATTTCAA 1080
QY 1081 TCTCGGCCAATAGGAGCGGATTTAAATACCTCAACGATGGGTCTACCAATACTTCTTATT 1140

DB 1081 TCTCGGCCAATAGGAGCGGATTTAAATACCTCAACGATGGGTCTACCAATACTTCTTATT 1140
QY 1141 AATCCTGTAGAGATTAATCACTTCTCTCGAGAGCTATATTGGACTGAATCATATGACAGGA 1200
DB 1141 AATCCTGTAGAGATTAATCACTTCTCTCGAGAGCTATATTGGACTGAATCATATGACAGGA 1200
QY 1201 GTGCTTCTATGGGGAATTTTACCTTTGAAAGAGGTACTGCTAACTATATAGTCAACCTATGAG 1320
DB 1201 GTGCTTCTATGGGGAATTTTACCTTTGAAAGAGGTACTGCTAACTATATAGTCAACCTATGAG 1320
QY 1261 TTTTAGGAAACCTCTCAGAAATACTTTTGAAGAGGTACTGCTAACTATATAGTCAACCTATGAG 1380
DB 1261 TTTTAGGAAACCTCTCAGAAATACTTTTGAAGAGGTACTGCTAACTATATAGTCAACCTATGAG 1380
QY 1321 TCACCTGGGCTTCAATTTAAAGATTTCAAGAACTGAAATTTACACGAGAAACAAACAGAACGA 1380
DB 1321 TCACCTGGGCTTCAATTTAAAGATTTCAAGAACTGAAATTTACACGAGAAACAAACAGAACGA 1380
QY 1381 CCAAAATATGAATCATATAGTCAATAGTATCTCAGATAGGGCTCAATTTCCAAATCTTAGG 1440
DB 1381 CCAAAATATGAATCATATAGTCAATAGTATCTCAGATAGGGCTCAATTTCCAAATCTTAGG 1440
QY 1441 GTGCATGTACCACTATATTTCTTGGAGCGCACCGTAGTGCAGATCGTACAAATACCATTAGT 1500
DB 1441 GTGCATGTACCACTATATTTCTTGGAGCGCACCGTAGTGCAGATCGTACAAATACCATTAGT 1500
QY 1501 TCAGATAGCATAACACAAATACCAATTTGGTAAATCAATTTCAACCTTAATTCAGGTACTCT 1560
DB 1501 TCAGATAGCATAACACAAATACCAATTTGGTAAATCAATTTCAACCTTAATTCAGGTACTCT 1560
QY 1561 GTAGTCAGTGGCCAGGATTTACAGAGGGGATATTAATCCGAACTAACGTTAAATGGTAGT 1620
DB 1561 GTAGTCAGTGGCCAGGATTTACAGAGGGGATATTAATCCGAACTAACGTTAAATGGTAGT 1620
QY 1621 GTACTAAGTATGGGTCTTAATTTTAAATATACATCAATTTACAGCGGTATCCGCTGAGAGTT 1680
DB 1621 GTACTAAGTATGGGTCTTAATTTTAAATATACATCAATTTACAGCGGTATCCGCTGAGAGTT 1680
QY 1681 CGTTATGCTGCTTCTCAACAAATGGTCTCGAGGGTAACTGTGCGAGGGAGTACTACTTTT 1740
DB 1681 CGTTATGCTGCTTCTCAACAAATGGTCTCGAGGGTAACTGTGCGAGGGAGTACTACTTTT 1740
QY 1741 GATCAAGGATTCCTCTAGTACTATGAGTGCATAATGAGTCTTTGACATCTCAATCATTTAGA 1800
DB 1741 GATCAAGGATTCCTCTAGTACTATGAGTGCATAATGAGTCTTTGACATCTCAATCATTTAGA 1800
QY 1801 TTTTGCAGAAATTTCTGTAGGTATTTAGTGCATCTGGCAGTCAAACTGCTGGAATAAGTATA 1860
DB 1801 TTTTGCAGAAATTTCTGTAGGTATTTAGTGCATCTGGCAGTCAAACTGCTGGAATAAGTATA 1860
QY 1861 AGTAATAATGCGGTAGACAAACGTTTCACTTTGATAAATTTGAATTCATTTCCAAATTTACT 1920
DB 1861 AGTAATAATGCGGTAGACAAACGTTTCACTTTGATAAATTTGAATTCATTTCCAAATTTACT 1920
QY 1921 GCAACCTTCCAGCAGATACAGATTTAGAAAGGGCGCAAGCGGTGAATGCTCTGTTT 1980
DB 1921 GCAACCTTCCAGCAGATACAGATTTAGAAAGGGCGCAAGCGGTGAATGCTCTGTTT 1980
QY 1981 ACTAATACGAATCCAAAGAGATTTGAAACAGATGTGACAGATTATCATATTTGATCAAGTA 2040
DB 1981 ACTAATACGAATCCAAAGAGATTTGAAACAGATGTGACAGATTATCATATTTGATCAAGTA 2040
QY 2041 TCCAAATTTAGTGGCGTATTTATCGGATGATTTCTGCTTAGATGAAAGAGAGAAATTTACTT 2100
DB 2041 TCCAAATTTAGTGGCGTATTTATCGGATGATTTCTGCTTAGATGAAAGAGAGAAATTTACTT 2100
QY 2101 GAGAAAGTGAATATGCGAAACGACTCAGTGTAGAAAGAACTTTACTCCAAAGATCCAAAC 2160
DB 2101 GAGAAAGTGAATATGCGAAACGACTCAGTGTAGAAAGAACTTTACTCCAAAGATCCAAAC 2160
QY 2161 TTTCAATCCATTAATAGCAACAGACTTCAATCTATTAATGAGCAATCGAATTTTCCACA 2220

[illegible]


```
Qy 2341 CCGACGTAATTTATATCAAAAAATAGAGAGTTCGGAATTTAAAGCTTATATCTCGCTACCAA 2400
Db 2341 CCGACGTAATTTATATCAAAAAATAGAGAGTTCGGAATTTAAAGCTTATATCTCGCTACCAA 2400
Qy 2401 TTAAGAGGTATATTGAAGATAGTCAAGATTAGAGATATATTTGATTCGTTAATAGCG 2460
Db 2401 TTAAGAGGTATATTGAAGATAGTCAAGATTAGAGATATATTTGATTCGTTAATAGCG 2460
Qy 2461 AAACATGAACATTGGATGTTCCAGGTACCGAGTCCGATCGCGCTTTTCAGTTGAAAGC 2520
Db 2461 AAACATGAACATTGGATGTTCCAGGTACCGAGTCCGATCGCGCTTTTCAGTTGAAAGC 2520
Qy 2521 CCAATCGAAGGTGCGAGAACCGAATCGATCGGCAACAATTTTGAATGGAATCCTGAT 2580
Db 2521 CCAATCGAAGGTGCGAGAACCGAATCGATCGGCAACAATTTTGAATGGAATCCTGAT 2580
Qy 2581 CTAGATTGTTCTTCACAGATGGAGAAAAATGCGCATCATCTCCCATCATTTCTCTTTG 2640
Db 2581 CTAGATTGTTCTTCACAGATGGAGAAAAATGCGCATCATCTCCCATCATTTCTCTTTG 2640
Qy 2641 GATATTGATATTGGATGCACAGACTTGCATGAGAAATCTAGGCGTGTGGGTGATTTCAAG 2700
Db 2641 GATATTGATATTGGATGCACAGACTTGCATGAGAAATCTAGGCGTGTGGGTGATTTCAAG 2700
Qy 2701 ATTAAGACGAGGAAGTCTATGCAAGACTAGGGAATCTGGAATTTATTTGAAGAGAAACCA 2760
Db 2701 ATTAAGACGAGGAAGTCTATGCAAGACTAGGGAATCTGGAATTTATTTGAAGAGAAACCA 2760
Qy 2761 TTATTAGGAGACACTGCTCTGCTGTCAGAGACGACAGAGAAAAATGGAGACAAACGT 2820
Db 2761 TTATTAGGAGACACTGCTCTGCTGTCAGAGACGACAGAGAAAAATGGAGACAAACGT 2820
Qy 2821 GAAAAAATCAATTTGAAAAAAGAGTATATACAGAGGCAAAAGAGCTGTGATGCT 2880
Db 2821 GAAAAAATCAATTTGAAAAAAGAGTATATACAGAGGCAAAAGAGCTGTGATGCT 2880
Qy 2881 TTATTGTAAGTCTCAATTAATATAGATTAACAGCGGATCAAAACATTTGGCATGATTCAT 2940
Db 2881 TTATTGTAAGTCTCAATTAATATAGATTAACAGCGGATCAAAACATTTGGCATGATTCAT 2940
Qy 2941 CGGCGACATAAATGTTTCATCGAATTCGAGAGGCTTATCTGTCAGAAATATCTGTTATC 3000
Db 2941 CGGCGACATAAATGTTTCATCGAATTCGAGAGGCTTATCTGTCAGAAATATCTGTTATC 3000
Qy 3001 CCGGCTGTAATGCGGAAATTTTGAAGATTAGAAGTTCGCAATTAATCACTGCAATCTCC 3060
Db 3001 CCGGCTGTAATGCGGAAATTTTGAAGATTAGAAGTTCGCAATTAATCACTGCAATCTCC 3060
Qy 3061 CTATACGATCGAGAAATGTCGTTTAAATATGTCATTTTAAATATGATGATGATGCTGG 3120
Db 3061 CTATACGATCGAGAAATGTCGTTTAAATATGTCATTTTAAATATGATGATGATGCTGG 3120
Qy 3121 AATGTAAGGCGATGTAGATGTACACAGAGCCATCACCGTTCCTGCTGTTATCCCA 3180
Db 3121 AATGTAAGGCGATGTAGATGTACACAGAGCCATCACCGTTCCTGCTGTTATCCCA 3180
Qy 3181 GAATGGGAAGCAGAAAGTGTCAAGCAGTTCGCGTCTGTCGGGGCGTGGCTATATCCTC 3240
Db 3181 GAATGGGAAGCAGAAAGTGTCAAGCAGTTCGCGTCTGTCGGGGCGTGGCTATATCCTC 3240
Qy 3241 CGTGTCAAGCGTCAAAAGAGGATATGGAGAGGTTGTGTAAACGATCCATGAAATCGAG 3300
Db 3241 CGTGTCAAGCGTCAAAAGAGGATATGGAGAGGTTGTGTAAACGATCCATGAAATCGAG 3300
Qy 3301 AACAAATCAGACGAATTAATTTAAAACTGTGAAGAGAGGAGTGTATCCAAACGAT 3360
Db 3301 AACAAATCAGACGAATTAATTTAAAACTGTGAAGAGAGGAGTGTATCCAAACGAT 3360
Qy 3361 ACAGAAACGTGTAATGATTAATCTGCAACCAAGTACAGCAGTATGTAATTTCCCGTAAT 3420
Db 3361 ACAGAAACGTGTAATGATTAATCTGCAACCAAGTACAGCAGTATGTAATTTCCCGTAAT 3420
```

```
Qy 3421 GCTGGATATGAGTGCATATGAAGTTGATCTACAGCATCTGTTAATTTACAAACCGACT 3480
Db 3421 GCTGGATATGAGTGCATATGAAGTTGATCTACAGCATCTGTTAATTTACAAACCGACT 3480
Qy 3481 TATGAAGAGAAACGTTATACAGATGTACGAAGAGATAATCATTTGTAATATGACAGAGGG 3540
Db 3481 TATGAAGAGAAACGTTATACAGATGTACGAAGAGATAATCATTTGTAATATGACAGAGGG 3540
Qy 3541 TATGTGAATTTATCCACATCTACAGCTGGTTATATGA CAAGAANAATTAGAAATATTTCCCA 3600
Db 3541 TATGTGAATTTATCCACATCTACAGCTGGTTATATGA CAAGAANAATTAGAAATATTTCCCA 3600
Qy 3601 GAAACCGATTAAGCTATGATTTGAGATTGGAAGAACGGAAGGAGTTTATTTGTAGACAGC 3660
Db 3601 GAAACCGATTAAGCTATGATTTGAGATTGGAAGAACGGAAGGAGTTTATTTGTAGACAGC 3660
Qy 3661 GTGGAATTAATCTCTTATGAGAGGAATAG 3687
Db 3661 ATAGAATTAATCTCTTATGGAAGATAG 3687

RESULT 3
US-10-428-961-62
; Sequence 62, Application US/10428961
; Publication No. US2003023711A1
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Wei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Rupar, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin
; FILE REFERENCE: Polynucleotides, Compositions, and Methods of Use (Amended)
; CURRENT APPLICATION NUMBER: US/10/428,961
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/661,322
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: 60/153,995
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 62
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-10-428-961-62

Query Match 88.4%; Score 3260.2; DB 17; Length 3684;
Best Local Similarity 93.3%; Pred. No. 0;
Matches 3445; Conservative 0; Mismatches 233; Indels 15; Gaps 3;

Qy 1 TTGACTTCAAAATAGGAAAAATGAAATGAAATTAATAATGCTTTATCGATTCCAGCTGTA 60
Db 1 TTGACTTCAAAATAGGAAAAATGAAATGAAATTAATAATGCTTTATCGATTCCAGCTGTA 60
Qy 61 TCGAATCATTTCCACAAAATGGATCTATCAACCGATGCTCGTATTTAGGATCTTTTGT 120
Db 61 TCGAATCATTTCCGACAAAATGAAATCTATCAACCGATGCTCGTATTTAGGATGCTTTGT 120
Qy 121 ATAGCCGAGGGGAATTAATATCAATCCACTTTGTTAGCCCATCAACAGTCCAAACGGTATT 180
Db 121 ATAGCCGAGGGGAACAAATATCGATCCATTTGTTAGCCCATCAACAGTCCAAACGGTATT 180
Qy 181 AACATAGCTGTAGAAATCTAGGTGTTATAGGCGTACCGTTTGTGACAAATAGTAGT 240
Db 181 AACATAGCTGTAGAAATCTAGGTGTTATAGGCGTACCGTTTGTGACAAATAGTAGT 240
Qy 241 TTTTATAGTTTTCTTGTGTTGTAATATGCCCCCGGCGAGAGATCATAGTGGGAAATTTTC 300
Db 241 TTTTATAGTTTTCTTGTGTTGTAATATGCCCCCGGCGAGAGATCATAGTGGGAAATTTTC 300
Qy 301 CTAGAACATGTGGAACAACTTTAATAATCAACAAATAACAGAAATGCTAGGAATACGGCA 360
```



```
QY 2515 GAAAGCCCAATCGAAGGTCGGAGAACCGAATCGATCGGCAACACATTTTGAATGGAAT 2574
DB 2512 GAAAGCCCAATCGAAGGTCGGAGAACCGAATCGATCGGCAACACATTTTGAATGGAAT 2571
QY 2575 CCTGATCTAGATTGTTCTCGCAGAGATGGAGAAAAATGTGGCATCATTCCTCCATCATTTTC 2634
DB 2572 CCTGATCTAGATTGTTCTCGCAGAGATGGAGAAAAATGTGGCATCATTCCTCCATCATTTTC 2631
QY 2635 TCTTTGGATATTGATATTGGATGCACAGACTTGCATCAGAAATCTAGCGGTGTGGGTGGA 2694
DB 2632 TCTTTGGATATTGATATTGGATGCATAGACTTGCATCAGAAATCTAGCGGTGTGGGTGGA 2691
QY 2695 TTCAGAGTTAAGACGCGAAGGTCATGCAAGACTAGGGAATCTGGAAATTTATTGAAGAG 2754
DB 2692 TTCAGAGTTAAGACGCGAAGGTCATGCAAGACTAGGGAATCTGGAAATTTATTGAAGAG 2751
QY 2755 AAACCATTTATTAGGAGAACGACTGCTCGTGTGAAGAGAGCAGAGAAAAAATGGAGAGAC 2814
DB 2752 AAACCATTTATTAGGAGAACGACTGCTCGTGTGAAGAGAGCAGAGAAAAAATGGAGAGAC 2811
QY 2815 AAACGCGTGAACAACTACAAATTTGGAACAAAAACGAGTATATACAGAGGCAAAAGAGCTGTG 2874
DB 2812 AAACGCGTGAACAACTACAAATTTGGAACAAAAACGAGTATATACAGAGGCAAAAGAGCTGTG 2871
QY 2875 GATGCTTTATTGTAGATTCTCAATATAATAGATTACAAGCGGATACAAACATTTGGCATG 2934
DB 2872 GATGCTTTATTGTAGATTCTCAATATATAGATTACAAGCGGATACAAACATTTGGCATG 2931
QY 2935 ATTTCATGCGCAGATAAATGTTTCATCGAATTCGAGAGGCTTATCTGTGAGAAATATCT 2994
DB 2932 ATTTCATGCGCAGATAAATGTTTCATCGAATTCGAGAGGCTTATCTTTGAGAAATATCT 2991
QY 2995 GTTATCCCGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAGGTCGCATTTACACTGCA 3054
DB 2992 GTTATCCCGGTGTAAATGCGGAAATTTTGAAGAAATTTAGAGGTCGCATTTACACTGCA 3051
QY 3055 ATCTCCCTTATACGATGCGAGAAATGTCGTTTAAAAATGGTGATTTTAAATATGGAATGCA 3114
DB 3052 ATCTCCCTTATACGATGCGAGAAATGTCGTTTAAAAATGGTGATTTTAAATATGGAATGCA 3111
QY 3115 TGTGGAATCTTAAAGGCGATGTAGATGTACACAGAGCCATCACCGTCTGTCTCTGTT 3174
DB 3112 TGTGGAATGTAAAGGCGATGTAGATGTACACAGAGCCATCACCGTCTGTCTCTGTT 3171
QY 3175 ATCCAGAAATGGGAAGCAGAGTGTCACAAAGCAGTTCGCGTCTGTCCGGGCGTGGCTAT 3234
DB 3172 ATCCAGAAATGGGAAGCAGAGTGTCACAAAGCAGTTCGCGTCTGTCCGGGCGTGGCTAT 3231
QY 3235 ATCTCCGTGTACAGGTCACAAAGAGGATATGAGAGGGTGTGTAAACGATCCATGAA 3294
DB 3232 ATCTCCGTGTACAGGTCACAAAGAGGATATGAGAGGGTGTGTAAACGATCCATGAA 3291
QY 3295 ATCCAGAAATACAGAGCACTTAAATTTTAAAACTGTGAGAGAGGAGTGTATCCA 3354
DB 3292 ATCCAGAAATACAGAGCACTTAAATTTTAAAACTGTGAGAGAGGAGTGTATCCA 3351
QY 3355 ACGGATACAGGAACGCTGTATGATTATCTACTGCAACCAAGGTACAGCAGTATGTAATTC 3414
DB 3352 ACGGATACAGGAACGCTGTATGATTATCTACTGCAACCAAGGTACAGCAGTATGTAATTC 3411
QY 3415 CGTAATCTCGATATGAGGATGCAATATGAAGTTGATACTACAGCATCTGTAAATACAAA 3474
DB 3412 CGTAATCTCGATATGAGGATGCAATATGAAGTTGATACTACAGCATCTGTAAATACAAA 3471
QY 3475 CCGACTTATGAGAGAGAAACGTATACAGATGTACGAGAGATATCTTGTGAATATGAC 3534
DB 3472 CCGACTTATGAGAGAGAAACGTATACAGATGTACGAGAGATATCTTGTGAATATGAC 3531
QY 3535 AGAGGGTATGTGAATTTTCCACCACTACAGCTGGTTATATGACAAAGAAATTAGAATAC 3594
DB 3532 AGAGGGTATGTGAATTTTCCACCACTACAGCTGGTTATATGACAAAGAAATTAGAATAC 3591
```

```
QY 3595 TTCCAGAAACCCGATAGGATGATGATTCGATTCGAGAAACCGGAGGAAAGTTTATTGTA 3654
DB 3592 TTCCAGAAACCCGATAGGATGATGATTCGATTCGAGAAACCGGAGGAAAGTTTATTGTA 3651
QY 3655 GACAGCGTGGAAATTTACTCTCTTATGGAGGAATAG 3687
DB 3652 GACAGCGTGGAAATTTACTCTCTTATGGAGGAATAG 3684

RESULT 4
US-10-809-953-9
; Sequence 9, Application US/10809953
; Publication No. US20040181825A1
; GENERAL INFORMATION:
; APPLICANT: Van Mellaert, Herman
; APPLICANT: Botterman, Johan
; APPLICANT: Van Rie, Jeroen
; APPLICANT: Joos, Henk
; TITLE OF INVENTION: RECOMBINANT PLANT EXPRESSING NON-COMPETITIVELY BINDING Bt INSECTIC
; TITLE OF INVENTION: CRYSTAL PROTEINS
; FILE REFERENCE: 021565-078
; CURRENT APPLICATION NUMBER: US/10/809,953
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: US/09/661,016
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: PCT/EP90/00905
; PRIOR FILING DATE: 1990-05-30
; PRIOR APPLICATION NUMBER: GB 89401499.2
; PRIOR FILING DATE: 1989-05-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 3687
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3687)
US-10-809-953-9

Query Match 85.0%; Score 3133.6; DB 18; Length 3687;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 3373; Conservative 0; Mismatches 299; Indels 30; Gaps 3;

QY 1 TTGACTTCAAAATAGGAAAAATGAGAAATGAAATATAAAATGCTTTATCGATTCCAGCTGTA 60
DB 1 TTGACTTCAAAATAGGAAAAATGAGAAATGAAATATAAAATGCTTTATGAGGATAGCTTGTGT 45
QY 61 TCGAATCATTTCCACACAAATGGATCTATCACCAGATGCTCGTATTGAGGATCTTTTGTGT 120
DB 46 TCGAATCATTTCCGACAAATGGATCTATCACCAGATGCTCGTATTGAGGATAGCTTGTGT 105
QY 121 ATAGCGGAGGGGAATATATCAATCCACTTTGTTAGCGCATCAACAGTCCAAACGGGTATT 180
DB 106 ATAGCGGAGGGGAACAAATATGATTCCTATTGTTAGCGCATCAACAGTCCAAACGGGTATT 165
QY 181 AACATAGCTGGTGAATATCTAGGTGTATTAGGCGTACCGTTGCTGACAAATAGTAGT 240
DB 166 AACATAGCTGGTGAATATCTAGGCGTATTGGGCGTACCGTTGCTGACCAACTAGTAGT 225
QY 241 TTTTATAGTTTTCTTGTGTGTAATTTATGCCCCCGCGGAGAGATCAGTGGGAAATTTTC 300
DB 226 TTTTATAGTTTTCTTGTGTGTAATTTATGCCCCCGCGGAGAGATCAGTGGGAAATTTTC 285
QY 301 CTAGAACATCTCGAACCAACTTATAAATCAACAAATACAGAAATGCTAGGAATACCGCA 360
DB 286 CTAGAACATCTCGAACCAACTTATAAATCAACAAATACAGAAATGCTAGGAATACCGCT 345
QY 361 CTTCGCTGATTACAAGGTTTATGAGGATTCCTTTAGAGCTTATCAACAGTCACTTTGAAGAT 420
DB 346 CTTCGCTGATTACAAGGTTTATGAGGATTCCTTCAGAGCTTATCAACAGTCACTTTGAAGAT 405
QY 421 TGGCTAGAAAAACCGTGTGATGATGCAAGAACGAGAAAGTGTCTTTTATATACCCCAATATATAGCC 480
```

Db 406 TGGCTAGAAAACCGTGATGATGCAAGAACGAGAAGTGTTCTTCATACCCAAATATATAGCT 465
 Qy 481 TTAGAACTTGATTTTCTTAATGCGATGCGCTTTTCGCAATTAAGAAACCAAGAGTTCCCA 540
 Db 466 TTAGAACTTGATTTTCTTAATGCGATGCGCTTTTCGCAATTAAGAAACCAAGAGTTCCCA 525
 Qy 541 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTATTGAGAGATGCTCT 600
 Db 526 TTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTATTATTATTGAGAGATGCTCT 585
 Qy 601 CTTTTCGTAGTGAATTTGGGCTTTACATCGCAGGAAATTCACGTTATTATTGAGCGCAA 660
 Db 586 CTTTTCGTAGTGAATTTGGGCTTTACATCGCAGGAAATTCACGCTATTATTGAGCGCAA 645
 Qy 661 GTGGAACAAACGAGAGATTTATCCGACTATTTGCGTAGAATGGTATATAACAGTCTAAAT 720
 Db 646 GTGGAACGAAACGAGAGATTTATCCGACTATTTGCGTAGAATGGTATATAACAGTCTAAAT 705
 Qy 721 AGCTTGAGAGGACAAATGCGCAAGTTGGTGCTTATTAATCAATTCGCTAGAGATCTA 780
 Db 706 AGCTTGAGAGGACAAATGCGCAAGTTGGTAGCGTATTAATCAATTCGCTAGAGATCTA 765
 Qy 781 ACGTTAGGGTATTAGATCTAGTGGCACTATTCCTCAAGCTATGACACTCGCACTTATCCA 840
 Db 766 ACGTTAGGATATTAGATCTAGTGGCACTATTCCTCAAGCTATGACACTCGCACTTATCCA 825
 Qy 841 ATAAATACGAGTCTCAGTTAAACAGGAGTTTATACAGCGCAATTTGAGCAACAGGG 900
 Db 826 ATAAATACGAGTCTCAGTTAAACAGGAGTTTATACAGCGCAATTTGAGCAACAGGG 885
 Qy 901 GTAATATGCAAGTATGAATGGTATTAATAAATGAACCTTCGTTTCGCTATATAGAG 960
 Db 886 GTAATATGCAAGTATGAATGGTATTAATAAATGAACCTTCGTTTCGCTATATAGAG 945
 Qy 961 ACTCGGTTATCCGAAGCGGATCTACTGATTTCTAGAAACCTTACAAATTTTATGAG 1020
 Db 946 GCTCGGCTATCCGAAGCGGATCTACTGATTTCTAGAAACCTTACAAATTTTATGAG 1005
 Qy 1021 ACTTCATCAGATGAGTGTCTAGGCAATATGATCTTACTGGGGGGGACACAAATTCAA 1080
 Db 1006 GCTTCATCAGATGAGTGTCTAGGCAATATGATCTTACTGGGGGGGACACAAATTCAA 1065
 Qy 1081 TCTCGGCAATFAGAGGCGGATTAATACTCAACGATGGGTCTACCAATATCTTCTATT 1140
 Db 1066 TCTCGGCAATFAGAGGCGGATTAATACTCAACGATGGGTCTACCAATATCTTCTATT 1125
 Qy 1141 AATCCTGTAGATTAATCAATCTTCTCGAGACGTATATTGGACTGAATCATATGCAAGGA 1200
 Db 1126 AATCCTGTAAATTAACGGTTTCGATCTCGAGACGTATATTGGACTGAATCATATGCAAGGA 1185
 Qy 1201 GTGCTTCTATGGGAATTTACCTTGAACCTTATTCATGCTGCTCCTACTGTTAGATTTAAT 1260
 Db 1186 GTGCTTCTATGGGAATTTACCTTGAACCTTATTCATGCTGCTCCTACTGTTAGATTTAAT 1245
 Qy 1261 TTTAGGAACCTCAGAAATCTTTTGAAGAGGTACTGCTAACTATAGTCAACCTATGAG 1320
 Db 1246 TTTAGGAACCTCAGAAATTTCTGTATAGAGGTACCGCTAACTATAGTCAACCTATGAG 1305
 Qy 1321 TCACCTGGCTTCAATTTAAAGATTAGAAACGTGAATTCACAGAAACCAACAGAAACGA 1380
 Db 1306 TCACCTGGCTTCAATTTAAAGATTAGAAACGTGAATTTACACAGAAACCAACAGAAACGA 1365
 Qy 1381 CCAAAATTAATGAATCATATAGTATAGTATCTCAGATAGGCTTCTTTCACAAATCTPAGG 1440
 Db 1366 CCAAAATTAATGAATCATATAGTATAGTATCTCAGATAGGCTTCTTTCACAAATCTPAGG 1425
 Qy 1441 GTGCATGTAACAGTATATCTTGGACCGAGTATGAGATCGTACAAATACCAATPAGT 1500
 Db 1426 GTGAATGTACCGGTATATCTTGGACCGATCTGTAGTGCAGATCGTACGAATACGATTTGGA 1485
 Qy 1501 TCAGATAGCATACACAAATACCAATTCGTAATTCATTCACCTTAAATTCAGGTACCTCT 1560

Db 1486 CCAATAGAAATCACCCAAATCCAAATGGTAAAGCATCCGAATCTCCTCAAGGTACCAC 1545
 Qy 1561 GTAGTCAAGTCCCGCAGGATTTACAGGAGGAGATATAATCCGAATCAAGTTTAATGGTAGT 1620
 Db 1546 GTTGTATAGAGGACGAGGATTTACTGGTGGGATATTCTTCGAAGCAACGAATACTGGTGA 1605
 Qy 1621 GTACTAAGTATGGTCTTTAATTTAATAATACATCATTTACAGCGGTATCGCTGAGAGTT 1680
 Db 1606 TTTGGACCGATAGAGTAATCTGTTAAACGACCAATTAACAACAGATATCGTATAGGATTC 1665
 Qy 1681 CGTTATGCTCTTTCTCAAAACAATGGTCTCGAGGGTAACTGTCGGAGGGAGTACTACTTTT 1740
 Db 1666 CGCTATGCTTCAACTGTAGATTTTGTATTTCTTTGTATCACGTGGAGTACTACTGTAAAT 1725
 Qy 1741 GATCAAGGATTCCTTAGTACTATGAGTCAAAATGAGTCTTTTGACATCTCAATCATTTAGA 1800
 Db 1726 AATTTTAGATTTCTTAGTACTATGAGTCAAAATGAGTCAAAATGAGTCAAAATGAGTCA 1785
 Qy 1801 TTTGCAAGATTTCTCTGTAGGTATTAGTGCATCTGGCAGTCAA--ACTGCTGGAATTAAGT 1857
 Db 1786 AGAGTGTCTTTTACTACACCTTTTACTTTTACACAAATTCAGATATAATTCGACGTCT 1845
 Qy 1858 ATAAGTAATTAATGCAGGTAGACAAAAGTTTCACTTTGATAAAATGGAATTCATTTCCAAT 1917
 Db 1846 ATTCAAGGCTTGTAGTGGAAATGGGAAAGTGTATATAGATAAAATTTGAAATTTATTCAG 1905
 Qy 1918 ACTGCAACCTTCGAACGAGATACGATTTAGAAAGGCGCAAGGCGGTGATGCTCTG 1977
 Db 1906 ACTGCAACCTTCGAACGAGATATGATTTAGAAAGGCGCAAGGCGGTGATGCTCTG 1965
 Qy 1978 TTTACTAATACGAATCCAAAGATTTGAAACAGATGTGACAGATTTATCATATTTGATCAA 2037
 Db 1966 TTTACTAATACGAATCCAAAGATTTGAAACAGATGTGACAGATTTATCATATTTGATCAA 2025
 Qy 2038 GTATCCAAATTTAGTGGGCTTTTATCGGATGAATTTCTGCTTAGATGAAAGAGAGATTA 2097
 Db 2026 GTATCCAAATTTAGTGGGCTTTTATCGGATGAATTTCTGCTTAGATGAAAGAGAGATTA 2085
 Qy 2098 CTTGAGAAAGTGAATAATGCGAAACGACTCAGTGTAGAAAGAACTTTACTCCAGATCCA 2157
 Db 2086 CTTGAGAAAGTGAATAATGCGAAACGACTCAGTGTAGAAAGAACTTTACTCCAGATCCA 2145
 Qy 2158 AACTTCACATCCATTAATAGCAACAGACTTTCATATCTAATAAGCAATCGAAATTC 2217
 Db 2146 AACTTCACATCCATTAATAGCAACAGACTTTCATATCTACTAATAGCAATCGAAATTC 2205
 Qy 2218 ACATCTATCCATGAAACATCTGAAACATGATGGTGGGAGAGTGAAGAACATTAACATCCAG 2277
 Db 2206 ACATCTATCCATGAAACATCTGAAACATGATGGTGGGAGTGAAGAACATTAACATCCAG 2265
 Qy 2278 GAAGGAAATGACGTATTTTAAAGAGAAATTAAGTGCACACTACCGGGGACTTTTAAATGAGTGT 2337
 Db 2266 GAAGGAAATGACGTATTTTAAAGAGAAATTAAGTGCACACTACCGGGGACTTTTAAATGAGTGT 2325
 Qy 2338 TATCCGACGTATTTATCAAAAATAGGAGAGTCCGAAATTTAAAGCTTATATCTCGCTAC 2397
 Db 2326 TATCCGACGTATTTATCAAAAATAGGAGAGTCCGAAATTTAAAGCTTATATCTCGCTAC 2385
 Qy 2398 CAATTAAGAGGTATATTGAAGTAGTCAAGATTTAGAGATATAATTTGATTCGTTATAAT 2457
 Db 2386 CAATTAAGAGGTATATTGAAGTAGTCAAGATTTAGAGATATAATTTGATTCGTTATAAT 2445
 Qy 2458 GCGAAACATGAACATTTGGATGTTCCAGGTACCGAGTCCGATGGCGCTTTTCAGTTGAA 2517
 Db 2446 GCGAAACATGAACATTTGGATGTTCCAGGTACCGAGTCCCTATGGCGCTTTTCAGTTGAA 2505
 Qy 2518 AGCCCAATCGAAGGTGGGAGAACCGAATCGATGCGCACCAATTTTGAATGGAATCCCT 2577
 Db 2506 AGCCCAATCGAAGGTGGGAGAACCGAATCGATGCGCACCAATTTTGAATGGAATCCCT 2565
 Qy 2578 GATCTAGATTTGTTCTTCGAGAGATGAGAAATGTCGATCATCTCCATCATTTCTCTCT 2637
 Db 2566 GATCTAGATTTGTTCTTCGAGAGATGAGAAATGTCGATCATCTCCATCATTTCTCTCT 2625

```

Qy 2638 TTGGATATTGATATTGGATGCACAGACTTGCATCAGAAATCTAGCGCTGTGGGTGGTATTC 2697
Db 2626 TTGGATATTGATATTGGATGCACAGACTTGCATCAGAAATCTAGCGCTGTGGGTGGTATTC 2685
Qy 2698 AAGATTAAAGCGCAGGAGGTCATCGAAGACTAGGAAATCTGGAAATTTATTGAGAGAAA 2757
Db 2686 AAGATTAAAGCGCAGGAGGTCATCGAAGACTAGGAAATCTGGAAATTTATTGAGAGAAA 2745
Qy 2758 CCATTATTAGAGAGCACTCTCTCGTGTGAAGAGCAGAGAGAGAGAGAGAGAGAGAGAG 2817
Db 2746 CCATTATTAGAGAGCACTCTCTCGTGTGAAGAGGCGAGAGAGAGAGAGAGAGAGAGAG 2805
Qy 2818 CGTGAAAACTCAATTTGGAAAAACAGAGTATATACAGAGGCAAAAGAGCTGTGGAT 2877
Db 2806 CGTGAAAACTCAATTTGGAAAAACAGAGTATATACAGAGGCAAAAGAGCTGTGGAT 2865
Qy 2878 GCTTTATTGTAGATTTCTCAATTAATAGATTACAAGGGATACAAACATTTGGCATGATT 2937
Db 2866 GCTTTATTGTAGATTTCTCAATTAATAGATTACAAGGGATACAAACATTTGGCATGATT 2925
Qy 2938 CATCGGCGAGATAAATTTGTTTCATCGAATTCGAGAGGCTTATCTGTGAGAAATTTATCTGTT 2997
Db 2926 CATCGGCGAGATAAATTTGTTTCATCGAATTCGAGAGGCTTATCTGTGAGAAATTTATCTGTT 2985
Qy 2998 ATCCGGGTGTAATGCGGAAATTTTGAAGAAATTTAGAGGTGCGATATACCTGCAATC 3057
Db 2986 ATCCAGGTGTAATGCGGAAATTTTGAAGAAATTTAGAGGTGCGATATACCTGCAATC 3045
Qy 3058 TCCCTATACGATGCGAGAAATGTCGTTAAATAATGGTGAATTTTAAATGATGATAGCATGC 3117
Db 3046 TCCCTATACGATGCGAGAAATGTCGTTAAATAATGGTGAATTTTAAATGATGATAGCATGT 3105
Qy 3118 TGGAAATGTAAGGCGATGTAGATGTACAACAGAGCCATACCGTCTGTCTGCTTTGTTATC 3177
Db 3106 TGGAAATGTAAGGCGATGTAGATGTACAACAGAGCCATACCGTCTGTCTGCTTTGTTATC 3165
Qy 3178 CCAGAAATGGGAAGCAGAAAGTGTCAACAGCACTTCGCGTCTGTCCGGGGCGTGGCTATATC 3237
Db 3166 CCAGAAATGGGAAGCAGAAAGTGTCAACAGCACTTCGCGTCTGTCCGGGGCGTGGCTATATC 3225
Qy 3238 CTCGCTGCACAGGTCAAAAGAGGATATGAGAGGGTGTGTGTAAACGATCCATGAATC 3297
Db 3226 CTCGCTGCACAGGTCAAAAGAGGATATGAGAGGGTGTGTGTAAACGATCCATGAATC 3285
Qy 3298 GAGAACCAATACAGACGAACTAAATTTTAAATACTGTGAAGAGAGGAGTGTATCCAAAG 3357
Db 3286 GAGAACCAATACAGACGAACTAAATTTTAAATACTGTGAAGAGAGGAGTGTATCCAAAG 3345
Qy 3358 GATACAGGAACGTGTATGATTATCTGACACCAAGGTACAGC-----AGTA 3405
Db 3346 GATACAGGAACGTGTATGATTATCTGACACCAAGGTACAGCTGGATGCGCAGATGCA 3405
Qy 3406 TGTAATTCCTGTAATGCTGGATATGAGATGATGATGATGATGATGATGATGATGATGATG 3465
Db 3406 TGTAATTCCTGTAATGCTGGATATGAGATGATGATGATGATGATGATGATGATGATGATG 3465
Qy 3466 AATTACAAACCGCTTATGAGAGAGAACGTATACAGATGTACAGAGAGATTAATCATGTT 3525
Db 3466 AATTACAAACCGCTTATGAGAGAGAACGTATACAGATGTAAAGAGAGATTAATCATGTT 3525
Qy 3526 GAATATGACAGAGGTATGTGAATTTATCCACCACTTACCAGCTGGTGTATATGACAAAAGAA 3585
Db 3526 GAATATGACAGAGGTATGTGAATTTATCCACCACTTACCAGCTGGTGTATGACAAAAGAA 3585
Qy 3586 TTAGAATATCTCCAGAAACCGATGATGATGATGATGATGATGATGATGATGATGATGATG 3645
Db 3586 TTAGAATATCTCCAGAAACCGATGATGATGATGATGATGATGATGATGATGATGATGATG 3645
Qy 3646 TTTATTGTAGACGCTGGAATTTACTCTTATGAGGAGATAG 3687
Db 3646 TTTATTGTAGATGCGTGGAAATTTACTCTCTCATGGAAGATAG 3687

```

```

RESULT 5
US-09-826-660-22
; Sequence 22, Application US/09826660
; Patent No. US20010026940A1
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Steilman, Steven J.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; FILE REFERENCE: MA-714XC2D1
; CURRENT APPLICATION NUMBER: US/09/826,660
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/178,252
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: 60/065,215
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/076,445
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 3558
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic B.t. toxin gene
US-09-826-660-22

```

```

Query Match 42.9%; Score 1581.6; DB 9; Length 3558;
Best Local Similarity 66.5%; Pred. No. 0;
Matches 2458; Conservative 0; Mismatches 1084; Indels 156; Gaps 7;

Qy 2 TGACTTCAAAATAGGAAAAATCAGAAATGAAATATATAATGCTTTATCGATTCAGCTGTAT 61
Db 2 TGACTTCTACAGAAAGAACGAGAACGAGATCATCAAGCTCTTTCTATCCAGCTGTT 61
Qy 62 CGAATCATTCACAAAATGGAATCTATCAACAGATGCTCGTATTGAGGATTCCTTTGTGA 121
Db 62 CTAACCATTTCTGCTCAGATGAACCTTTCTACTGATGTAGAAATCGAGGATTCCTTTGCA 121
Qy 122 TAGCGGAGGGAATTAATCAATCCACTTGTAGCGCATCAACAGTCCAAACGGTATTA 181
Db 122 TTGCTGAGGGAACCAACATTCATTCCTGCTTCTACTGTTCAAACTGGAATCA 181
Qy 182 ACATAGCTGTAGAAATCTAGGTGATTTAGCGGTACCGTTTGTCTGGACAAATAGCTAGT 241
Db 182 ACATGCTGGAAGATCTCTGGAGTTCCTGGAGTTCCTGCTGGACAGATTCCTTCTT 241
Qy 242 TTTATAGCTTTTCTTTGTTGGTGAATTTATGGCCCGCGCAGAGATCAGTGGGAAATTTTCC 301
Db 242 TCTACTCTTTCTTTGTTGGAGAGCTTTTGGCTAGGGGAAGAGATCCTTGGGAGATCTTCC 301
Qy 302 TAGAACATGTCGAACAACTTATAAATCAACAAATACAGAAATGCTAGGAATACGGCAC 361
Db 302 TTGAGCATGTGAGCAGTTGATTCGTCAACAAAGTTACTTGAGAACACTAGAGATCTGCTC 361
Qy 362 TTGCTCCATACAGGTTTATAGGATTTCTTTAGAGCTATCAACAGTCACTTGAAGATT 421
Db 362 TTGCTAGACTTCAAGGACTTTGGAACCTTTTCAGAGCTTACCAACATCTCTTGAGGATT 421
Qy 422 GGCTAGAAAAACCGTGTAGTATGAGAACGAGAGAGTGTCTTTATACCAATATATAGCCT 481
Db 422 GGCTTGAGAACAGAGATGATGCTAGAACTAGATCTGTGTGTTGTAACACTCAGTACATGCTC 481
Qy 482 TAGAATCTGATTTCTTAATGCGATGCGCTTTTTCGCAATTTAGAAAACCAAGATTCAT 541
Db 482 TTGAGCTTGATCTTCTGAAACGCTATGCAATTTGCTATCAGAAAACCAAGAGTTCAC 541
Qy 542 TATTAAATGGTATATGCTCAAGCTGCAAAATTTTACACCTTATTATTATGAGAGATGCTCTC 601
Db 542 TTCTCATGGGTGATGCTCAAGCTGCTAACCTTCTTCTTCTTTAGAGATGCTAGCT 601

```


Qy 602 TTTTGGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACGTTATTATGAGCGCAAG 661
 Db 602 TGTTCGGATCTGAGTTCGGACTTACTTCTCAAGAGATTCAGGATCTACGAGAGCAAG 661
 Qy 662 TGGAAACGAGAGATTTATCCGACTATTCGCTAGATGATGATTAATACAGGTTCAATA 721
 Db 662 TTGAGAGACTAGAGAGTACTCTGACTACTGCGCTAGGTGATCAACACCTGGAATGAACA 721
 Qy 722 GCTTTGAGAGGCAAAATCGCCAAAGTTGGGTCGCTTATAATCAATTCGCTAGAGATCTAA 781
 Db 722 ACCTTAGAGGAACCTAAACGCTGAGTCTTGGCTTAGATACCAACAGTTGAGAGATCTTA 781
 Qy 782 CGTTAGGGGTATTAGATCTAGTGCACTATTTCCAAAGTATGACATCGCGACTTATCCAA 841
 Db 782 CTCTTGGAGTTCTTGATCTTGTGCTTGTTCCTTGTTCCTTGTTCCTTGTTCCTTGTTC 841
 Qy 842 TAAATACGAGTGTCTCAGTTAACAGGAGGTTTATATACAGCGCAATTTGGAGCACAGGG 901
 Db 842 TGAACACTTCTGCTCAACTTACTAGAGAGATCTACATCTGATCCAAATCGGAAGAACTAAC 901
 Qy 902 TAAATATGCGCAAGTATGAATTTGGTATATAATTAATGCACTTTCGTTTTCGCTTA 955
 Db 902 CTCATCTGGATTCGCTTCTACTTAATCTGTTTCAACACACGCTTCAATCTTCTCTGCTA 961
 Qy 956 TAGAGACTGCGGTTATCCGAAGCCGCACTACTCTGATTTCTTAGAACAACTTACAAATTT 1015
 Db 962 TCGAGGCTGAGTGCATCAGACCAACCACTCTTCTTGACTTCCAGAGCAACTTACTATCT 1021
 Qy 1016 TTAGCACTTCAACAGATGAGTGTCTACTAGGCATATGACTTACTGCGGGGGGACACAA 1075
 Db 1022 TCTCTGTTCTTTTCTAGATGCTTAACACTCAGTACATGAACTACTGCGGTTGGACATAGAC 1081
 Qy 1076 TTCAATCTCGGCCCAATAGGAGCGGATTAATACCTCAAGCGATGGTCTACCAATACCT 1135
 Db 1082 TTGAGTCTAGAACTATCAGAGGATCTCTTCTACTTCTACTCTAGTAAACACTTACACT 1141
 Qy 1136 CTATTAATCTCTGAAGATTAATCTTCTCTCGAGACGTATATGCACTGAACTCAATATG 1195
 Db 1142 CTATCAACCCAGTTACTCTTCAGTTCTTCTAGAGATGTTGACAGAACTGAGCTTTTCG 1201
 Qy 1196 CAGAGTGTCTTCTAGGGGATTTACCTTGAACCTATTCATGTTGTCCTTACTGTTAGAT 1255
 Db 1202 CTGGAATCAACAT-----TCTTCTTACTACTCCAGTGAACGAGTTCTCTGGGCTAGAT 1255
 Qy 1256 TTAATTTTAGGAACCTCAGAACTTTTGAAGAGGTACTGCTTAACATAAGTCAACCCCT 1315
 Db 1256 TCAACTGGAGAAACCCATGAACTCTCTT--AGAGGTTCTTGTGTACACATTTGGAT 1312
 Qy 1316 ATGAGTCACTGGGCTTCAATTAAGATTAAGATCAGAACTGAAATACCAACCAAGAACAG 1375
 Db 1313 ACACCTGGAGTTGGTACCCAGTTGTTGATTTCTGAGACTGAGCTTCCACAGAGACTACTG 1372
 Qy 1376 AACGACCAATATGAATCATATAGTTCATAGTTATCTCATAGGCTCATTTTCAAT 1435
 Db 1373 AGAGCAAACTACGAGTCTTACTCTCATAGACTTCTTAACATCTGTTTGTATCTCTGGAA 1432
 Qy 1436 CTAGGTTGCATGTACACAGTATTTCTTGGAGCAACCGTAGTGAGATCGTACAAATACCA 1495
 Db 1433 ACACCTTAGAGCTCCAGTGTACTCTTGGACTCATAGATCTGCTGATAGAACTAACACCA 1492
 Qy 1496 TTAGTTCAGATAGATTAACAAATACCAATTTGGTAAATCATTCACCTTAAATTCAGGTA 1555
 Db 1493 TCTCTTCTGATCTTATCACTCAGATTCACCTTGTGAAGTCTTCAACTTGAACCTCTGGAA 1552
 Qy 1556 CCTCTGAGTCAAGTCCGAGGATTTACAGAGGGATATAATCCGAACTAAAGTTAATG 1615
 Db 1553 CTTCTGTTGTTCTTGGACCAAGGATTTACTGGAGAGACATCATCAGAACTAAAGTGAAC 1612
 Qy 1616 GTAGTGTACTAAGTATGGGCTTAAATTTAATATATCATCATACAGCGGTATCGCGTGA 1675
 Db 1613 GATCTGTTCTTCTATGAGGATTTGAACTTCAACACACTTCTCTTCAAGATACAGAGTTA 1672
 Qy 1676 GAGTTCGTTATGCTGCTTCTCAACAAATGTTCTGAGGGTAACTGTCGGAGGGAGTACTA 1735

Db 1673 GAGTTAGATACGCTGCTTCTCAAACTATGTTCTTAGAGTTACTGTTGGAGGATCTACTA 1732
 Qy 1736 CTTTTCGATCAAGGATTTCCCTAGTACTATGATGCAATGATGCTTTGACATCTCAATCAT 1795
 Db 1733 CTTTTCGATCAAGGATTTCCCATCTACTATGCTGCTAACGAGTCTCTTACTTCTCAATCTT 1792
 Qy 1796 TTAGATTGTCAGAAATTTCTGTAGGTATAGTGCATCTGCGAGTCAAACTGCTGGAAATA 1855
 Db 1793 TCAGATTGCTGAGTTCCCAAGTTGGAATCTCTGCTTCTGATCTCAAACTGCTGGAATCT 1852
 Qy 1856 GTATAGTATAATGAGGATGAGCAAAAGTTTCACTTTTGTATATAATTTGAATTCATTTCAA 1915
 Db 1853 CTATCTCTTAAACAGCTGGAAGACAACTTTCCACTTCCGACAAAGATTTGAGTTCAATTTCAA 1912
 Qy 1916 TTACTGCAACTTTCGAAGCAGAAATAGATTTAGAAAGGCGCAAGGCGGTGAAATGCTC 1975
 Db 1913 TCACTGCTCTCTCGAGGCGAGTCTGACTTTGGAAGAGCAGAGAGCGGTGAAATGCTC 1972
 Qy 1976 TGTTTACTTAATACGAATCCAAAGAGATTTGAAACAGATGTGACAGATTTATCATATTGATC 2035
 Db 1973 TGTTCATCTCGTCCAATCAGATTGGGCTCAAGACAGATGTGACTGACTATCACATCGATC 2032
 Qy 2036 AAGTATCCAAATTTAGTGGCGTGTATTCGGATGAAATCTGCTTAGATGAAAGAGAGAAAT 2095
 Db 2033 GCGTTTCCAACTTGTGAGTGCCTCTCTGATGAGTTCTGTTGGATGAGAGAGAGAGAT 2092
 Qy 2096 TACTTCGAAGAGTGAATATCGAAACGACTGATGATGAAGAACTTACTTCCAGATC 2155
 Db 2093 TGTCCGAGAGGTCAAACATGCTTAAGCGCTTAGTGATGAGCGGAACTTGTCTTCAAGATC 2152
 Qy 2156 CAAACTTTCATCATCCATCAATTAAGCAACAGACTTCTATCTACTAATGAGCAATCGAAT 2215
 Db 2153 CCACTTTCCGGGATCAACAGGCACTA----- 2181
 Qy 2216 TCATCATCTATCCATGAACAACTCTGAACATGGATGGTGGGAGAGTGAAGACATTTACAAATCC 2275
 Db 2182 -----GATCGTGGATGGAGGGAGAGTACGGACATCACCATTC 2218
 Qy 2276 AGGAAGAAATGACGTATTTAAAGAGAAATTTACGTACACTTACCGGGGACTTTTAAATGAT 2335
 Db 2219 AAGAGTGTATGATGTTTCAAGGAGAACTATGTTAGCTCTTTGGGTACCTTTGATGATG 2278
 Qy 2336 GTTATCCGAGCTATTTATATCAAAAAATAGGAGGTCCGAAATTTAAAGCTTTATCTCGCT 2395
 Db 2279 GCTATCCAACTACCTGTACCAAGATAGATGAATCGAACTCAAAGCTCACACAGAT 2338
 Qy 2396 ACCAATTAAGAGGTATTTGAAGATAGTCAAGATTTAGAGATATATTTGATTCGTTATTA 2455
 Db 2339 ACCAGTTGAGAGGTATCATCGAGGACAGTCAAGACCTTGAGATCTACCTCATCAGATACA 2398
 Qy 2456 ATCGAAACATGAACATTTGGATGTTCCAGGTACCGATCCGATCCGCTTTTCAGTTG 2515
 Db 2399 ACCCAACATGAGACAGTCAATGTGCTTGGAGCGGTTCATCTCGGCCACTTTTACGCC 2458
 Qy 2516 AAAGCCCAATCGGAAGGTGCGGAGAACCGAATCGATTCGCAACCACTTTTGAATGGAATC 2575
 Db 2459 CAACTCCCATCG----- 2471
 Qy 2576 CTGATCTAGATTGTTCTCTGAGAGATGGAGAAATAATGTGGGATCATTTCCCATCTTCT 2635
 Db 2472 -----CAAGTGTGCCCATCATCTCACACCTTCT 2500
 Qy 2636 CTTTGGATATTGATATTTGGATGCACAGCTTGCATGAGAAATCTAGCGGTGTTGGTGTAT 2695
 Db 2501 CTTTGGACATAGACGTTGGCTGTACCGACTGAAACGACCTCCGTTGTGGTGTATCT 2560
 Qy 2696 TCAAGATTGAAGCAGCAGGAGGTCTATGCAAGCTAGGAACTCTGGAATTTTATTTGAAGAGA 2755
 Db 2561 TCAAGATCAAGACTCAAGATGGCCATGCCAGGCTAGGCAATCTGGAGTTTCTAGAGAGAGA 2620
 Qy 2756 AACCATTTAGAGAGAGCACTGCTCTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA 2815


```

Db 2621 AACACCTTTGGTGAAGAGCCCTGCTAGAGTGTAAGAGGGCTGAGAAGAGTGGAGGACA 2680
Qy 2816 AACGTGAAATACAAATTTGAAACAAACAGAGTATATACAGAGGCAAAAGAGCTCTGG 2875
Db 2681 AGAGAGAGAGTTGGAATGGGAAACAAACATTTGTACAAAGAGCCAAAGAGCGCTG 2740
Qy 2876 ATGCTTTATTTGTAGATCTCAATATATATAGATTACAGGGGATACAAACATTTGGCATGA 2935
Db 2741 ACGCTCTGTTGTGAACCTCAGTATGATAGGCTCCAAGCTGATACCAATAGCTATGA 2800
Qy 2936 TTCAATCGGCGAGATAAAATCTTTCATCGAATTCGAGAGGCTTATCTGTCAAGAAATTAATCTG 2995
Db 2801 TTCAATCGTGCAGACAAACGGCTTCATAGCAATTCGGAGAGCTTACCTTCTCGAATTCAGC 2860
Qy 2996 TTATCCCGGCTGTAATTCGCGAAATTTTGAAGAAATTTAGAGGTGCGAATTTATCACTGCAA 3055
Db 2861 TGATTCGGGCTGCAATGCTATCTTTGAAGAGTTAGAAGGCGCATCTTCACTGCAT 2920
Qy 3056 TCTCCCTATAGGATCGAGAAATGCTGTTAAATGTTGATTTTAATATGATAGCAT 3115
Db 2921 TCTCCTTGTATGATCGGAGGAATGTCATCAAGATGTTGATCTTCAACAAATGGCCTATCCT 2980
Qy 3116 GCTGGAATGTAAGGCGCATGATCT---ACAACAGAGCCATCACCGTTCTGTCCTTG 3172
Db 2981 GCTGGAATGTAAGGCGCATGATGTAAGAGAACAGAAATCACCGCTCTGTCCTTG 3040
Qy 3173 TTATCCAGAAATGGGAAGCAGAGTGTCAACAGCAGTTCGCGTCTGTCGCGGCGTGGCT 3232
Db 3041 TTGTTCTGATGGGAAGCAGAGTTCACAGAAAGTTTCTGCTGCTGCTGCTGGCT 3100
Qy 3233 ATATCTCCGTGTACAGGTACAAAGAGGATATGAGAGGTTGTTGTAACGATCCATG 3292
Db 3101 ACATTTCTGTTTACCGCTACAAAGAGGATACGAGAGGTTGCGTCACCATACAG 3160
Qy 3293 AATTCGAGAACTACAGAGCACTAAATTTAAAACTGTGAAGAGAGGAGTGTATC 3352
Db 3161 AGATTGAGAAACACACGAGCTGAAGTTACAGCAACTGCGTGCAGAGGAAAGTCTACC 3220
Qy 3353 CAACGGATACAGGAACGTGTAATGATTATCTGACACCAAG-----GTACAGCAGTAT 3406
Db 3221 CAACACACCGTAACTTGCATGACTACACTGGACTCAAGAGGAGTATGAGGTTACTT 3280
Qy 3407 GTAATTCCTGTAATGCTGGATATGAGGATGATATGAAGTTGATATACAGCATCTGTTA 3466
Db 3281 ACATTTCTCGAATCGAGGATACGATGGAGCTATGAGAGCAACTCTTCTGTACCCGCTG 3340
Qy 3467 ATTACAAACCGACTTATGAGAGAAACGATATACAGATGTACGAGAGATATCATTTGTG 3526
Db 3341 ACTATGATCAGCCTATGAGGAGAGGCTTACACCGATGAGCGTAGGGAACAATCTTTGCG 3400
Qy 3527 AATATGACAGAGGATGTGAATTTATCCACCTACCCAGCTGGTTATATGACAAAAAGAT 3586
Db 3401 AATTAACAGAGCTATGGGACTACACACGTTACACCGGCTATGTCACCAAGAGT 3460
Qy 3587 TAGAATCTTCCGAAACCGATAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 3646
Db 3461 TAGAGTACTTTCCGAAACCCCAAGGTTTGGATTGAGATTGAGAAACCGAAGGAACAT 3520
Qy 3647 TTATTTGACAGCGGTGGAATTAATCTTATGAGGAA 3684
Db 3521 TCATTGTTGATAGCGTGGAGTTACTTCTGATGGAGAA 3558

```

RESULT 6

```

US-09-988-462-6
; Sequence 6, Application US/09988462
; Publication No. US20030046726A1
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; Desai, Nalini M.
; Lewis, Kelly S.
; Kramer, Vance C.
; Warren, Gregory W.

```

```

; Evola, Stephen V.
; Crossland, Lyle D.
; Wright, Martha S.
; Merlin, Ellis J.
; Launis, Karen L.
;
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; INSECTICIDAL ACTIVITY IN MAIZE
;
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Syngenta Biotechnology, Inc.
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/988,462
; FILING DATE: 20-No. US20030046726A1-2001
; CLASSIFICATION DATA: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/547,422
; FILING DATE: 11-APR-2000
; APPLICATION NUMBER: US 08/459,504
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951,715
; FILING DATE: 25-SEP-1992
; APPLICATION NUMBER: US 07/772,027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-188051
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
;
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; optimized cry1B"
; /note= "disclosed in Figure 6."
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-988-462-6

```

```

Query Match 40.6%; Score 1495.6; DB 10; Length 3624;
Best Local Similarity 63.7%; Pred. No. 0;
Matches 2310; Conservative 0; Mismatches 1299; Indels 15; Gaps 2;

Qy 79 ATGGATCTATCACAGATGCTCGTATTGAGGATCTTTTGTATAGCCGAGGGAATAAT 138
Db 1 ATGGACCTGTGCGCCGACGCCGATCGAGGACAGCTGTGCGCCGAGGCAACAAC 60
Qy 139 ATCAATCCACTTGTAGCGCATCAACAGTCCAAACCGGTTATTAACTAGTGTGTAATA 198
Db 61 ATCGACCCCTTCTGTAGCGCAGCACCGTGCAGACGGCATCAACATCGCCGCGCATC 120
Qy 199 CTAGGTGTTATAGCGGTACCGTTTCTGGCAAAATAGCTAGTTTATAGTTTCTTGT 258
Db 121 CTGGGCGTGTGGGCGGTGCCCTTTCGCGGCGCAGCTGSCCAGCTTCTACAGCTTCTCCTG 180

```

Qy 259 GGTGAAATTATGGCCCGGCGAGAGATCAGTGGGAAATTTTCTTAGAACATGTGGAACA 318
 Db 181 GCGAGCTGTGGCCCGGCGCGCGGACGAGTGGAGATCTTCTTGGAGACGTGGAGCAG 240
 Qy 319 CTTATAAATCAACAATAAACAAGAAATGCTAGGAATACGGCACTTGTCTGATTAACAAGGT 378
 Db 241 CTGATCAACAGCAGATCACCGAGAACCGCCGCAACACCGCCCTGGCCCGCTGCAGGGC 300
 Qy 379 TTAGGAGATTCTTTAGAGCCTATCAACAGTCACTTGAAGATTGGCTAGAAAACCGTGAT 438
 Db 301 CTGGCGACAGCTTCTCGCGCTTACCAGCAGAGCCTGGAGGACTGGCTGGAGAACCGGAC 360
 Qy 439 GATCAAGAACGAGAAGTGTCTTTATACCAATATATAGCTTTAGAACTTGAATTTCTT 498
 Db 361 GACCCCGACCCCGAGCGTCTGTACACCCAGTACATCGCCCTGGAGCTGGAATCTCTG 420
 Qy 499 AATCGGATGCGGCTTTTCGCAATTAGAAACCAAGAGTTCATTTAATAGGTATATGCT 558
 Db 421 AACCCATGCCCTGTTCGCCATCCGCAACAGAGGTGCCCTGTCTGATGGTACGCC 480
 Qy 559 CAAGCTGGAATTTACACCTATTAATTTAGAGATGCTCTCTTTTGGTAGTAATTT 618
 Db 481 CAGCCGCCAACCTTGCACCTGCTGCTGCGGACGCCAGCCTGTTTCGGCAGCGAGTTC 540
 Qy 619 GGGCTTACATCGCAGGAAATTCACAGTATTATGAGCGCCAGTGGAAACAAACGAGAGAT 678
 Db 541 GGCCTGACCCAGAGAGATCACCGCTACTACCGGCTGAAACAGCCTGCGCGCACCAAC 600
 Qy 679 TATTCGCACTATTCGAGAGATGATTAATAACAGTCTAAATAGCTTTGAGAGGACAAAT 738
 Db 601 TACAGCGACTACTGCGTGGAGTGATCAACAACCGGCTGAAACAGCCTGCGCGCACCAAC 660
 Qy 739 GCGCAAGTTGGGTGCGTTAATCAATTCGTTAGAGATCTAAGTTAGGGGTATTAGAT 798
 Db 661 GCGCCAGCTGGGTGCGCTCAACACAGTTCGCGCGACCTGACCCCTGGGCGTGTGGAC 720
 Qy 799 CTAGTGCACTATTCCCAAGCTATGACATCGCACTTATCAATAAATAACAGTGTCTCAG 858
 Db 721 CTGGTGGCCCTGTTCCTCAGCTACGACACCGGCACTTACCCATCAACACAGCGGCCAG 780
 Qy 859 TTAACAAGGGAAGTTTATACAGCGCAATTTGAGCAACAGGGGTAAATATGGCAAGTATG 918
 Db 781 CTGACCCGCGAGGTGTACCGAGCGCATCGGCGCCACCGCGTGAACATGGCCAGCATG 840
 Qy 919 AATTGGTATAATAATGACCTTCCTTTCGCTATAGAGACTGCGGTATTCGGAAGC 978
 Db 841 AACTGGTACAAACAACCGCCCGCAGCTTCAGCGCCATCGAGGCGCGCCATCCGCGAC 900
 Qy 979 CCGCATCTACTTGTATTTCTAGAACAACTTACAAATTTTATGACCTTCAATCAAGTGGAGT 1038
 Db 901 CCCACCTGTGGACTTCTGGAGCAGCTGACATCTTCAGCGCCAGCAGCGCCCTGGAGC 960
 Qy 1039 GCTACTAGGCATATGACTTACTGGCGGGGACAAATTCAAATCTCGGCCAATAGGAGGC 1098
 Db 961 AACACCCGCGACATGACTACTGGCGGGGACACCATCCAGAGCGCCGCCATCGCGGCG 1020
 Qy 1099 GGATTAATACCTCAAGCATGGGTCTACCAATCTTCTATTAATCTCTGTAAGATTATCA 1158
 Db 1021 GGCCTGAACACCAAGCACCGCGGCAACCAACACAGCATCAACCCCGTGACCCCTGCGC 1080
 Qy 1159 TTTCTTCTCGAGACGTATATGAGCTGAATCATATGAGAGTGTCTTATGGGGAATT 1218
 Db 1081 TTGCGCAGCGGAGAGTATACCGCACCGAGGCTACGCGCGGTGTCTGTGGGGATC 1140
 Qy 1219 TACTTGAACCTATTCATGTTGTCCTTACTGTTAGATTAAATTTAGGAACCCCTCAGAAT 1278
 Db 1141 TACCTGAGGCCATCCACGGGCTGCCACCGTGGCTTCAACTTCACCAACCCCGCAAGAC 1200
 Qy 1279 ACTTTTGAAGAGTACTGCTAATATAGTCAACCCCTATGATCACTGGCTTCAATTA 1338
 Db 1201 ATCAGCACCGCGGCAACCGCAACTACAGCCAGCCCTACGAGAGCCCTCGCGCTGCAGCTG 1260

Qy 1339 AAGATTTCAGAACTGAATTTACCACAGAAACACAGAACGACCAAAATTTATGATCATAT 1398
 Db 1261 AAGCAGCAGGACCGAGCTGCCCGCCGAGACACACCGAGCGCCCACTACGAGAGCTAC 1320
 Qy 1399 AGTCATAGGTTATCTCATATAGGCTCATTTCAAACTTAGGGGTGCATGTACCAAGTATAT 1458
 Db 1321 AGCCACCGCTGAGCCACATCGGCATCATCTGCGAGAGCGCGTGAACGCTGCCGCTGAC 1380
 Qy 1459 TCTTGGACGACCGTAGTCAGATCGTACAAATACCAATTAGTTTCAGATAGATACACAA 1518
 Db 1381 AGCTGGACCCACCGAGCGCGCCGACCAACACCAATCGGCCCAACCCGATCACCCAG 1440
 Qy 1519 ATACCATTTGTGTAATCATTTCAACCTTTAATTCAGGTACCTCTGTAGTCAGTGGGCCAGGA 1578
 Db 1441 ATCCCATGTTGAAGGCGAGGAGCTGCCCGAGGACCAACCGTGTGCGGGCCCGCGC 1500
 Qy 1579 TTTTACAGGAGGGGATATAATCCGAATTAACGTTAATGTTAGTGTACTTAAGTATGGGTCTT 1638
 Db 1501 TTCACGGCGCGACATCTCTGCGCCGACCAACACCGCGGCTTGGGCCCAATCCGCGTG 1560
 Qy 1639 AATTTTAATTAATACATTTACAGCGGTATCGCGTGAAGTTCGTTATGCTGTTCTCAA 1698
 Db 1561 ACCGTGAACGCGCCCTTGACCCAGCGCTACCGCATCGGCTTCCGCTACGCCAGCACCGTG 1620
 Qy 1699 ACAATGTCCTGAGGGTAACTGTGAGGAGTACTATTTTGATCAAGGATTCCTTAGT 1758
 Db 1621 GACTTCGACTTCTTGTGAGCGCGCGGCGCACACCGTGAACAACTTCCGCTTCTCGCG 1680
 Qy 1759 ACTATG---AGTGGAAATAGTCTTTGACATCTCAATCATTTAGATTGTCAGAAATTTCT 1815
 Db 1681 ACCATGAACAGCGCGACGAGCTGAAAGTACGGCAACTTCGTGGCGCGCGCTTCACCAACC 1740
 Qy 1816 GTAGGTATTAGTCACTGCGAGTCAAACTGCTGGATTAAGTAAATTAATGCAAGT 1875
 Db 1741 CCCTTACCTTACCAGATCCAGGACATCATCCGACACGATCCAGGCGCTGAGCGGC 1800
 Qy 1876 AGACAAACGTTTCACATTTGATAAAATTTGAATTTCAATTCCAATTTCTGCAACCTTCGAAGCA 1935
 Db 1801 AACGGCGAGGTGTACATCGACAGATCGAGATCATCCCGTGACCGGCCACCTTCGAGGCC 1860
 Qy 1936 GAATACGATTGAAGAGGGCGAAGAGCGCGGTGAATGTCTGTGTTTACTAATACGAATCCA 1995
 Db 1861 GAGTACGACCTGGAGCGCGCCAGAGCGCGTGAACGCGCTGTTTCAACAACAGCAACCCC 1920
 Qy 1996 AGAAGATTGAACACAGATGTGACAGATTATCATATTGATCAAGTATCCAAATTTAGTGGCG 2055
 Db 1921 CGCGCGCTGAAGACCGACGAGTGAACGACTACCATCGACCGGTGAGCAACCTTGGTGGCC 1980
 Qy 2056 TGTTTATCGGATGAATTTCTGCTTAGATGAAAAGAGAGAAATTTCTTGAGAAAGTGAATAT 2115
 Db 1981 TGCTGAGCGACGAGTTCTGCTGGACGAGAGCGCGAGCTGCTGGAGAGGTGAAGTAC 2040
 Qy 2116 GCGAAACGACTCAGTGATGAAGAAACTTACTCGAAGATCCAACTTCAATCCATCCATCAAT 2175
 Db 2041 GCCAAGCGCTGAGCGACGAGCGCAACTCTGCTGAGGACCCCACTTTCACGAGCATCAAC 2100
 Qy 2176 AAGCAACGACTTCTATCTACTAATGAGCAATCGAATTTTCACTATCTCCATGACAA 2235
 Db 2101 AAGCGCCGACTTCTATCAGCACCAAGCAGAGCAACTTCCACGACTCCAGAGCAG 2160
 Qy 2236 TCTGAACATGATGCTGGGGAAGTGAACATTTACAAATCCAGGAAGAAATGACGTATTT 2295
 Db 2161 AGCGAGCAGCGCTGGTGGGCGAGAGAAATCATCCATCCAGGAGGCAACGAGCTGTC 2220
 Qy 2296 AAGAGAAATTACGTACACTACCGGGGACTTTTAAATGAGTGTATTCGAGCTATTATAT 2355
 Db 2221 AAGGAGAACTTACGTACCTTCCCGCGCACCTTCAACGAGTGTACCCCACTTCTCTGAC 2280
 Qy 2356 CAAAAAATAGGAGTGGAAATTTAAAGCTTATCTGCTGCTACCAATTAAGAGGGGTATATT 2415
 Db 2281 CAGAGATCGCGAGAGCGAGCTGAGGCGCTACACCGCTACCGCTGCGCGGCTACATC 2340
 Qy 2416 GAAGATAGTCAAGATTTAGAGATATATTTTGATTGTTTAAATGCGAAACATGAACAATTG 2475

Qy	705	TAATACAGCGTCTAAATATAGCTTTGAGAGGGACAAATGCCCGAGATTGGGTGCGTTATAATCA	764
Db	627	CAATCAAGGATTTAGAAAATTTAAGAGGTACTAATACTCGAAATGGGCAAGATTTCAATCA	686
Qy	765	ATTCGGTAGAGATCTAAACGTTAGGGGTATATAGATCTAGTGGCACTATTTCCCAAGCTATGA	824
Db	697	GTTTAGGAGAGATTTAACACTTACGTGTATAGATATCGTTCTCTTTTTTCCGAACCTACGA	746
Qy	825	CACTCGCACTTATCCAAATAAATACGAGTGCTCAGTTAAACAAGGGAAGTTTATACAGACGC	884
Db	747	TGTTAGAACATATCCAAATTCAAACGTATCCCAATTAACAAGGGAATTTATACAGTTTC	806
Qy	895	AATTGGAGCAACAGGGGTAAATATATGCAAGTATGAATTTGGTATATAATAATATGCAACCTTC	944
Db	807	AGTAATTGAGGATTTCTCCAGTTTCTGC-----TAATATACCTAA	845
Qy	945	GTTTTCCGCTATAGAGACTGCGGTTATCCGAAGCCCGCATCTACTTGTATTTTCTTAGACA	1004
Db	846	TGGTTTAAATAGGCGGAATTTGGAAGTTAGACCGCCCATCTTATGGACTTTATGAA---	902
Qy	1005	ACTTACAAATTTTAGCACTTTCATCACGATGGAGTGCTACTAGGCATATGACTTACTGCGC	1064
Db	903	-----TTCCTTGTGTGTAACTGACAGACTGTTAGAGTCAAACTGTGTGGGAGG	953
Qy	1065	GGGGCACAAATTCAAATCTCGGCCAAATAGGAGGGGATTAATAACCTCAACGCATGGGTC	1124
Db	954	ACACTTAGTTAGTTACGAAATACGGCTGGTAAACCGTATAAATTTCCCTAGTTACGGGT	1013
Qy	1125	TACCAATACTTCTATTAATCTGTGAAGTATCATTTCTTCTCGAGACGTATATATGGAC	1184
Db	1014	CTTCAATCTCGGTGGCGGCATTTGGATTTGCAGATGAGGATCCAGCTCCTTTTATCGGAC	1073
Qy	1185	TGAATCATATCAGGAGTGCTTCTATGGGGAATTTACCTTCAACCTATTTCAATGTGTCCC	1244
Db	1074	ATTATCAGATCCTGTTTTGTCCGAGGAGT-----TTGGGAATCCTCAATATGTA	1127
Qy	1245	TACTGTTAGATTTAATTTTAGGAACCTTCAGAAATACTTTTGAAAGAGGTACTGCTAACTA	1304
Db	1128	GGGGCTTAGGGGAGTAGCATTTCAACAAACTGGTACGAACACACCCGAACATTTTAGAAA	1187
Qy	1305	TAGTCAACCTATGAGTCACCTGGGCTTCAATTAAGAATTCAGAACTGAATTAACACC	1364
Db	1188	TAGTGGGACCAATAGATTTCTCTAGATGAATCCCACTCAGGATTAATAGTGGGGACCC--TT	1246
Qy	1365	AGAAACACAGAACGACCAAAATTTAATCATATATAGTCATAGGTTTATCTCACATAGGCT	1424
Db	1247	GGATGATTATAGTCATGTATTAATCATGTTTACATTTGACGATGCCAGGTGAGATTT	1306
Qy	1425	CATTTCAAAATCTAGGGTGCAATGACAGTATATTTCTTTGGACGACCGTAGTGCAGATCG	1484
Db	1307	CAGGAAGTGATTCATGGAG--AGCTCCAATGTTTTCTTTGGACGCAACCGTAGTGCACCCC	1364
Qy	1485	TACAAATACCAATTAGTTCAGATAGCATTAACAATAACCATTTGGTAAATCATTTCAACCT	1544
Db	1365	TACAAATACAAATTCATCCGAGAGGATTTACTCNAATACCAATTTGGTAAAGCACATACCT	1424
Qy	1545	TAAATTCAGGTACTCTGTAGTCAGTGCCCGCAGGATTTACAGGAGGGGATATAATCCGAAC	1604
Db	1425	TCAGTCAGGTACTACTGTTGTAAAGGGCCCGGTTTACGGGAGAGATATCTTTCGACG	1484
Qy	1605	TAACTGTTAATCGGTAGTGTACTAAGTATGGGCTTAAATTTTAAATAATCATCATTTACAGC	1664
Db	1485	AACAAGTGGAGGACCATTTGCTTATACTATTTGTTAATAATAAATGGGCAATTAACCCAAAG	1544
Qy	1665	GTATCGGTGAGAGTTGCTTATGCTGCTTCTCAAAACAAATGGTCTCTGAGGTTAACTGTGCG	1724
Db	1545	GTATCGTGCAGAAATACGCTATGCTCTACTACAAATCTAAGAAATTTACGTAAACCGGTTC	1604
Qy	1725	AGGAGTACTACTTTTGATCAAGGATTCCTCTAGTACTATGATGCAATGACTCTTTGAC	1784
Db	1605	AGGTGAACGGAATTTTTCGTGCTCAATTTTAACAAAACAAATGATGATCCGTTGACCCATTAAC	1664
Qy	1785	ATCTCAATCATTTAGATTTGAGAAATTTTCTGTAGGTATTTAGTGCATCTGCGCATCAAC	1844

1665	Db	ATTCCAATCTTTTAGTTAGTTACGCAACTATTAAATACAGCTTTTACATTCCTCAATGAGCCAGAG	1724
1845	Qy	TGCT- --GGAATAAGTATAAGTAATAATCGAGGTAGACAAACGGTTTCACCTTTTGATAAAAT	1901
1725	Db	TAGTTTACAGTAGGTGCTGATCTTTTAGTTAGGGAATGAAGTTTATATATAGACAGATT	1784
1902	Qy	TGAATTCAATTCCAATTACTGCAACCTTTCGAAGCAGAAATACGATTTAGAAAGGGCGCAAGA	1961
1785	Db	TGAATTGATCCAGTTACTGCAACATTTGAAGCAGAAATATGATTTAGAAAGAGCACAAA	1844
1962	Qy	GGCGGTGAATGCTCTGTTTACTTAATACGAATCCAAGAAAGATTGAAACAGATGTGACAGA	2021
1845	Db	GGCGGTGAATGCGTGTGTTACTTCTATAAACCCAATAGGGATAAAAACAGATGTGACGGA	1904
2022	Qy	TTATCATATTGATCAAGATPATCCAAATTTAGTGGCGCTGTTATCGATCAATTTCTGTTAGA	2081
1905	Db	TTATCATATTGATCAGATATCCAAATTTAGTGGATTTGTTATCAGATGAAATTTTCTCTGGA	1964
2082	Qy	TGAAAAGAGAGAAATTACTTGGAGAAAGTGAATATGCCAAACGACTCAGTGATGAAGAAA	2141
1965	Db	TGAAAAGCAGAAATTGTCCGAGAAAGTCAAAACATGCCAAGCGACTCAGTGATGAGCGGAA	2024
2142	Qy	CTTACTCCAGATCCAAACTTTCACATCCATCCATCAATTAAGCAACCGAGACTTCATATCTACTAA	2201
2025	Db	TTTACTTTCAAGATCCAAACTTTCAAAGCGATCAATATAGGCAACTAGAC-----	2070
2202	Qy	TGACCAATCGAAATTTTCACATCTCATCCATGAACAATCTGAACATGGATGGTGGGGAAGTGA	2261
2071	Db	-----CGTGGTTGGAGAGGAAGTAC	2090
2262	Qy	GAAATTTACAAATCCAGGAAGAAATGACGTATTTAAAGAGAAATTCAGTCACTACCGGG	2321
2091	Db	GGATATTTACCATCCAAAGAGGAGTACGTTATTCAAAGAAAATTTATGTCACACTACCGAG	2150
2322	Qy	GACTTTTAATGAGGTGTTATCCGACGTATTTATATCAAAAAATAGGAGAGTCGGAATTTAAA	2381
2151	Db	TACCTTTGATGAGTGTCTATCCACGTATTTATATCAAAAAATAGTGAAGTCGAAAATTTAAA	2210
2382	Qy	AGCTTATATCTGCTACCAATTTAGAGGGTATATTTGAAGATAGTCAAGATTTAGAGATATA	2441
2211	Db	ACCTATACTCGTTATCAATTTAAGAGGGTATATCGAGGATAGTCAAGCTTTAGAAATCTA	2270
2442	Qy	TTTGATTCGTATAATCCGAAACATGAAACATTTGGATGTTCCAGGTACCGAGTCGGTATG	2501
2271	Db	TTTGATCCGCTATAATGCAAAACAGTAAATGCTGAGTACGGTCTCTTTATG	2330
2502	Qy	GCGCTTTCTAGTTGAAAGCCCAATTCGGAAGGTGCGGAGAAACCGAATCGATGCGCACCA	2561
2331	Db	GCGCTTTCTAGTCCAAAGTCCAAATCAGAAAAGTGTGGAGAAACCGAATCGATCGCGCCACA	2390
2562	Qy	TTTTGGAATGGAATCTTGATCTAGATTTGTTCTCCAGAGATGGAGAAAATGTCGCATCA	2621
2391	Db	CTTTGGAATGGAATCTTGATCTAGATTTGTTCTCCAGAGACGGGAAAATGTCGCATCA	2450
2622	Qy	TTCCCATCATTTCTCTTTGGATATTTGATATTTGGATGCACAGACTTGCATGAGAAATCTAGG	2681
2451	Db	TTCCCATCATTTCTCTTTGGACATTTGATTTGGATGTACAGACTTAATTCAGGACTTTAGA	2510
2682	Qy	CGTGTGGTGGTATTTCAAGATTTAAGACGCAAGGAAGTTCATGCAAGACTAGGGAAATCTGGA	2741
2511	Db	TGTATGGTGTATTTCAAGATTTAAGACGCAAGATGGCCATGCAAGACTAGGAAATCTTAGA	2570
2742	Qy	ATTTATTTGAGAGAAACCATTTATTAGGAGAGCACCTCTCTCGTGTGAAGAGACGACAGAA	2801
2571	Db	GTTTCTCGAAGAGAAACCATTTAGTCGGGGAAGCACATGACTCGTGTGAAAAGAGACAGAA	2630
2802	Qy	AAAATGAGAGACAAACGCTGAAAACCTACAAATTTGGAAAACAAAACGAGTATATACAGAGGC	2861
2631	Db	AAAATGAGAGATAAAGCTGAAAATTTGGAATTTGGAACAAAATATTGTTTATAAGAGGC	2690
2862	Qy	AAAAGAAGCTGTGGATGCTTTATTTGTAGATTTCTCAATATAAATAGATTTACAAGCGGATAC	2921

Db 2691 AAAGAAATCTAGATGCTTTATTTGTAAACTCTCAATATATGATCAATTTACAAGCGGATAC 2750
Qy 2922 AAACATTTGGCATGATTAATGCGGAGATAAATCTTTTCATCGAATTCGAGAGCTTATCT 2981
Db 2751 GAATATTTGCCATGATTCATGCGGAGATAAATCGTTTCATAGAAATTCGGGAAGCGTATCT 2810
Qy 2982 GTCAAGATTAATCTGTTATCCCGGGGTAAATCCGGAATTTTCAAGAAATAGAGGTG 3041
Db 2811 TCCAGAGTTATCTGTGATTCGCGGGGTAAATGTAGACATTTTCGAAGAAATTTAAAAGGCG 2870
Qy 3042 CATATCACTCAATCTCCCTATACGATCGGAGAAATGCTTTAAATATGCTGATTTTAA 3101
Db 2871 TATTTTCACTGCAATCTCTTATATGATCGGAGAAATGCTTTAAATATGCTGATTTTAA 2930
Qy 3102 TAATGGATTAATGCTGGAATTAAGAGGCGATGTAGATGT---ACAACAGAGCCATCA 3158
Db 2931 TAATGGCTTATCATGCTGGAACGTGGAAGGCGATGTAGATGTAGAGAAACAACCA 2990
Qy 3159 CCGTCTCTGCTGTTATCCAGAAATGGGAGAGAGATGTCAACAGCATTCGCGTCTG 3218
Db 2991 CCGTTCGGTCTGTTGTTCCGSAATGGGAGAGAGATGTCAACAGAGATTCGTTGCTG 3050
Qy 3219 TCCGGGGCGTGTATATCTCTGTCACAGCGTACAAAGAGGATATGAGAGGTTG 3278
Db 3051 TCCGGGTGCTGCTATATCTCTGTCACAGCGTACAAAGAGGATATGAGAGGTTG 3110
Qy 3279 TGTAAACATCATGAATTCGAGAAACAATACAGAGCAACTAAATTTTAAACCTGTGAAGA 3338
Db 3111 CGTAACCATTCATGAGATCGAGAAACAATACAGAGCAACTGAAGTTTAGCAACTCGGTAGA 3170
Qy 3339 AGAGGAAGTGTATCCACGGATACAGGAACTGTGTAATGATTAATCTGCAACACCAAGGTAC 3398
Db 3171 AGAGGAAGTGTATCCAAACCAACAGCGTAAACGTGTAATGATTAATCTGCAAAATCAAGAAGA 3230
Qy 3399 A-----GCAGTATGTAATCCCGTAACTGCTGGAATGAGATGCATATGAAGTTGATAC 3452
Db 3231 ATACGGGGTGGTACACTTCCCGTAACTGCTGGAATGAGATGCATATGAAGTTGATAC 3290
Qy 3453 TACAGCATCTGTTAATTAACAACCGACTTATGAAGAAAGAAACGATATACAGATGTACGAAG 3512
Db 3291 TTCTGTACCAAGCTGATATGCTGAGTATGAGAAAGAAATCGTATACAGATGTACGAAG 3350
Qy 3513 AGATAATCATTTGGAATATGACAGAGGATGTGAAATTTCCACCACTACAGCTGTTA 3572
Db 3351 AGACAAATCTTGTGAATTAACAGAGGATATGGGATTTACACCACTACAGCTGCTA 3410
Qy 3573 TATCACAAGAAATTAAGAACTTCCAGAAACCGATTAAGATGATGAGATTTGAGATTTGAGA 3632
Db 3411 TGTGACAAAGAAATTAAGATTAATCCAGAAACCGATTAAGATGATGAGATTTGAGATTTGAGA 3470
Qy 3633 AACGGAAGGGAAGTTTATTTGTAGACAGCGTGAATTTACTCTTATGAGGAA 3684
Db 3471 AACGGAAGGAACTTCTGTTGACAGCGTGAATTTACTCTTATGAGGAA 3522

RESULT 8
US-09-837-961-7
; Sequence 7, Application US/09837961
; Publication No. US20040058860A1
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel
; APPLICANT: Sick, August
; TITLE OF INVENTION: No. US20040058860A1el Bacillus thuringiensis Isolate Active Again
; TITLE OF INVENTION: and Genes Encoding No. US20040058860A1el Lepidopteran-Active Tox
; FILE REFERENCE: MA-43CDF2D3
; CURRENT APPLICATION NUMBER: US/09/837,961
; CURRENT FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/521,344
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: US 08/933,891
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: US 08/356,034
; PRIOR FILING DATE: 1994-12-14

; PRIOR APPLICATION NUMBER: US 08/210,110
; PRIOR FILING DATE: 1994-03-17
; PRIOR APPLICATION NUMBER: US 07/865,168
; PRIOR FILING DATE: 1992-04-09
; PRIOR APPLICATION NUMBER: US 07/451,261
; PRIOR FILING DATE: 1989-12-14
; PRIOR APPLICATION NUMBER: US 371,955
; PRIOR FILING DATE: 1989-06-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 3522
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
; US-09-837-961-7

Query Match 38.2%; Score 1408.4; DB 11; Length 3522;
Best Local Similarity 65.7%; Pred. No. 0;
Matches 2243; Conservative 0; Mismatches 1061; Indels 108; Gaps 9;

Qy 285 TCAGTGGGAAATTTTCTTAGAACATGTCGAACCACTTTATAAATCAACAAATAACAGAAAA 344
Db 207 TGATTGGAGCTTATTTCTTTTACAGATTGAACAATTGATTGAGCAAGAATAGAAACATT 266
Qy 345 TGCTAGGAATACGGCACTTCTGCTGATTACAGAGTTTAGGAGATTCCTTTAGAGCCTATCA 404
Db 267 GGAAGGAACCGGCAATTTACTACATTTACGAGGTTAGCAGATAGCTATGAAATTTATAT 326
Qy 405 ACAGTCACTTTGAGATTTGGCTAGAAACCGTATGATGCAAGACGAGAGTGTCTTTA 464
Db 327 TGAAGCACTAAGAGAGTGGGAAGCAATCTTAATAATGCACAATTAAGGGAAGATGTGCG 386
Qy 465 TACCAATATATAGCCCTTAGAACCTTGTATTTCTTAATCGCATGCGCTTTTCGCAATTTAG 524
Db 387 TATTCGATTTGCTAATACAGACGAGCTTTAATAACAGCAATAAATAATTTTACACTTAC 446
Qy 525 AAACCAAGAGTTCCATTTAATATGATGATGCTCAAGCTGCAAAATTTACACCTATTATT 584
Db 447 AAGTTTTGAAATCCCTTTTATCGGTCTATGTTCAAGCGCGCAATTTACATTTATCACT 506
Qy 585 ATTGAGAGATGCTCTCTTTTGGTGTAGTGAATTTGGGCTTACATCGCAGGAAATTCACG 644
Db 507 ATTAAGAGACGCTGTATCGTTTGGCAGGGTGGGAGCTGGATATAGCTACTGTTAATAA 566
Qy 645 TTATTTAGCGGCCAACAGTGGAAACAAACGAGAGATTTATTCGACTTATTCGCTAGAAATGGA 704
Db 567 TCATTATAATAGATTAAATAATCTTTATCATAGATATACGAAACATTTGTTGGACATA 626
Qy 705 TAATAAGGCTCTAATAAGCTTTGAGAGGGAACAATGCGCAAGTTGGGTGCGTTAATAATCA 764
Db 627 CAATCAAGGATTAAGAAACCTTAAGAGGTACTAATACTCGCAATGGGCAAGATTCATCA 686
Qy 765 ATTCGCTAGAGATCTACGTTTAGGGTATTAGATCTAGTGGCACTATTCGCCAGCTATGA 824
Db 687 GTTTAGAGAGATTTTAACTTACTGTATTAGATATCGTTGCTCTTTTTCGGAACCTAGA 746
Qy 825 CACTCGCACTTTTCCAAATAATACGAGTGTCTAGTTAAACAGGGAAGTTTATACAGACG 884
Db 747 TGTTAGAACATATCCAAATTCAAACGTCATCCCAATTAACAGGGAATTTATACAGTTC 806
Qy 885 AATTGGAGCAACAGGGGTAATAATATGGCAAGTATGAATTTGGTATATAATTAATATGACCTTC 944
Db 807 AGTAATTTGAGGATTCCTCAGTTCTGCTG-----TAATATACCTAA 845
Qy 945 GTTTTCGCTATAGAGACTCGGTTATCCGAAGCCCGCATCTACTTCAATTTTCTAGAAC 1004
Db 846 TGGTTTTTAATAGGCGGAATTTGGAGTTAGACCGCCCATCTTTATGGACTTTTATGAA--- 902
Qy 1005 ACTTACAATTTTGTAGCACTTCATCAAGTGGAGTGTACTAGGCATATGACTTACTTGGCG 1064
Db 903 -----TTCTTTGTTGTTAACTGCAAGACTGTTAGAGTCAAACTGTGGGAGG 953
Qy 1065 GGGGCAACAATTCATCTCGGCCAAATAGGAGCGGATTAATACTCAACGCATGGGTC 1124

Db 954 ACCTTAGTTACGAAATACGGCTGGTAACCGTATAAATTTCCCTAGTTACGGGGT 1013
 Qy 1125 TACCAATACCTTCTAATTAATCTGTAAGTATATCAATCTCTCTCGAGAGCTATATTGGAC 1184
 Db 1014 CTTCAATCTGGTGGCCCAATTTGGATTGTCAGATGAGGATCCACGTCCTTTTATACGGAC 1073
 Qy 1185 TGAATCATATACAGAGTCTTCTATATGGGAAATTTACCTTGAACCTATTCAATGTGTCCC 1244
 Db 1074 ATTATCAGATCTGTTTGTTCGAGGAGAT-----TTGGGAATCCTCATTAATGACT 1127
 Qy 1245 TACTGTTAGATTAAATTTAGGAACCTCAGAAATACATTTTGAAGAGGTACTGCTAACTA 1304
 Db 1128 GGGCTTAGGGGAGTAGCATTTCAACAACTGGTAGCAACACACCCGAAATTTAGAAA 1187
 Qy 1305 TAGTCAACCTTATGAGTACCTGGCTTCAATTAAGAGATTCAGAACTGAATTAACACC 1364
 Db 1188 TAGTGGGACCATAGATTCTCTAGATGAATCCACCTCAGGATTAATAGTGGGGCACC-TT 1246
 Qy 1365 AGAAACACAGAACGACCAAAATTAATGATCATATAGTCATAGGTTATCTCACATAGGGCT 1424
 Db 1247 GGAATGATTATAGTCATGTATTAATCATGTTACATTTGTACGATGGCCAGGTGAGATT 1306
 Qy 1425 CATTTCAAACTTAGGGTGCATGTACCAGTATATTTCTTGGACGACCCGTAGTCAGATCG 1484
 Db 1307 CAGGAAGTGATTCATGGAG--AGCTCCAAATGTTTCTTGGACGCACCGTAGTCAACCCC 1364
 Qy 1485 TACAAATACCAATTAGTTCAGATACATTAACAATACCAATGGTAAATCAATTCACCT 1544
 Db 1365 TACAAATACCAATTTGATCGGAGAGGATTAATCAATACCAATGGTAAAGACACATACCT 1424
 Qy 1545 TAAATTCAGGTACTCTGTAGTCAGTGGCCAGGATTTACAGGAGGAGATTAATTCGCAAC 1604
 Db 1425 TCAGTCAGGTACTACTGTTGTAAGAGGCCGGGTTTACGGGAGAGATATTTCTTCAGC 1484
 Qy 1605 TAACTGTTAATGGTAGTACTTAAGTATGGGTCTTAATTTTAATTAATACATTCATACAGC 1664
 Db 1485 AACAAAGTGGAGACCAATTTGCTTATCTATTAATGTTTAAATAAATGGGCAATTAACCCCAAAG 1544
 Qy 1665 GTATCGGTGAGAGTCTGTTATGCTGCTTCTCAACAAATGGTCTGAGGTTACTGTCGG 1724
 Db 1545 GTATCGGTGAGAGTACGCTATGCTCTACTACAAATCTAAGAAATTTACGTAACGGTTGC 1604
 Qy 1725 AGGAGTACTACTTTTGTATCAAGGATTCCTGTAGTACTATGATGCAATGAGTCTTTGAC 1784
 Db 1605 AGGTGAACGGATTTTGTCTGTCAATTTACAAACAAATGGATACCGGTGACCAATTAAC 1664
 Qy 1785 ATCTCAATCATTTAGATTGTCAGAAATTTCTGTAGTATTAGTGCATCTGGCAGTCAAC 1844
 Db 1665 ATTCCAATCTTTTAGTTACGCAACTAATTAATACAGCTTTTACATTCGCAATGAGCCAGAG 1724
 Qy 1845 TGCT--GGAAATGATTAATGATTAATGAGTACGAGTACGAAACGTTTCACTTTGATTAAT 1901
 Db 1725 TAGTTTACAGTAGGTGCTGATCTTTTAGTTACGGAATGAAGTTTATATAGACAGATT 1784
 Qy 1902 TGAATTCATTTCCAAATTAATGCAACTTCGAAGCAGAAATACGATTTAGAAAGGGCGCAAGA 1961
 Db 1785 TGAATTCATTTCCAGTTACTGCAATTTGAGCAGAAATGATTTAGAAAGAGCACAATA 1844
 Qy 1962 GGGGTTGAATGCTCTGTTTACTTAATACGAATTCGAAGAGATTGAAACAGATGTGACAGA 2021
 Db 1845 GGGGTTGAATGCTGTTTACTTCTATAAACCMAATAGGGATTAACAGATGTGACGGA 1904
 Qy 2022 TTATCATATGATCAAGTATCCAAATTTAGTGGCGGTGTTTATCGGATGAATTCGCTTAGA 2081
 Db 1905 TTATCATATGATCAAGTATCCAAATTTAGTGGATGTTTATCAGATGAATTTTGTCTGGA 1964
 Qy 2082 TGAAGAGAGAAATTAATTCGAGAAAGTCAAAATATGCAAAACGACTCAGTGTATGAAAGAAA 2141
 Db 1965 TGAAGAGAGAAATTTGTCGAGAAAGTCAAAATATGCAAAACGACTCAGTGTATGAGCGGAA 2024
 Qy 2142 CTTACTCAAGATCCAAATTCATCCATCAATTAAGCAACAGACTTCATATCTACTAA 2201

2025 TTTACTTCAAGATCCAAACTTTCAAGGCACTCAATAGGCAACTAGAC----- 2070
 Qy 2202 TGACCAATCGAATTTACATCTATCCATGAACAACTCTGAACATCGATGTTGGGGAAGTGA 2261
 Db 2071 -----CGTGGTTGGAGAGGAAGTAC 2090
 Qy 2262 GAACATTTACCAATCCAGGAAGGAATGACGTATTTTAAAGAGAAATTTACCTCACTACCGGG 2321
 Db 2091 GGATTTACCAATCCAAAGAGAGATGACGTATTTCAAGAGAAATTTATCTCACATACCAGG 2150
 Qy 2322 GACTTTTAATGAGTGTATCCGACGTATTTATATCAAAAAATAGGAGATCGGAATTTAAA 2381
 Db 2151 TACCTTTGATGAGTGTATCCAAAGTATTTATATCAAAAAATAGATGAGTCGAAATTTAAA 2210
 Qy 2382 AGCTTATACCTCGCTACCAATTTAGAGGGTATTTAGAGATAGTCAGATTTAGAGATATA 2441
 Db 2211 ACCCTATCTCGTTATCAATTAAGAGGGTATATCGAGGATAGTCAAGACTTTAGAAATCTA 2270
 Qy 2442 TTTGATTCGTTTAAATGCGAAACATGAACATTTGGATGTTTCCAGGTACCGGATCCGATG 2501
 Db 2271 TTTGATCCGCTATATGCAAAAACAGAAACAGTAAATGTCTAGGTACGGTCTTTATG 2330
 Qy 2502 GCCCTTTAGTTGAAGCCCAATCGGAAGTGGCGGAGAACCGAATCGATCGGACCAACA 2561
 Db 2331 GCCGCTTTAGTCCAAAGTCCAAATCAGAAAGTGTGGAGAACCGAATCGATCGCGGCCACA 2390
 Qy 2562 TTTTGAATGAATCTCGATCTAGATTTGTTCTCCAGAGATGGAGAAATAATGTCGCATCA 2621
 Db 2391 CTTTGAATGAATCTCGATCTAGATTTGTTCTCCAGAGACGGGGAATAATGTCGCATCA 2450
 Qy 2622 TTCCCATCATTTCTCTTTTGGATTTGATATTGGATGCAAGACTTGCATCAGAAATCTAGG 2681
 Db 2451 TTGCATCATTTCTCTCTTGGACATTTGATGTTGGATGTCAGACTTAAATGAGGACTTGA 2510
 Qy 2682 CGTGTGGTGGTATTTCAAGATTAAGACGAGGAGGTCATGCAAGACTAGGGAATCTGGA 2741
 Db 2511 TGTATGGTGGTATTTCAAGATTAAGACGAGGATGGCCATGCAAGACTAGGAAATCTAGA 2570
 Qy 2742 ATTTATTTGAAGAAACCAATTTAGGAGAACACTGTCGCTGTGAAGAGAGCAGAGAA 2801
 Db 2571 GTTTCTGAAAGAAACCAATTTAGTCCGGGAGACACTAGCTCTGTTGAAAGAGCAGAGAA 2630
 Qy 2802 AAAATGGAGAGACAAACGTTGAAACCACTCAATTTGAAACCAAAACGAGTATATACAGAGC 2861
 Db 2631 AAAATGGAGAGATAAACGTTGAAACCAATTTGAAATTTGAAACCAAAATATTGTTTATAAAGAGC 2690
 Qy 2862 AAAAGAGCTGTGGATGCTTTTATTTAGATTTCTCAATATAATAGATTACAGCGGATAC 2921
 Db 2691 AAAAGAAATCTGTAGATGCTTTTATTTGTAACCTCTCAATATGATCAATTTACAGCGGATAC 2750
 Qy 2922 AAACATTTGGCATGATTCATCGGCGAGATAAATCTGTTTTCATCGAAATTCGAGAGGCTTATCT 2981
 Db 2751 GAATTTGCCATGATTTATCGGCGAGATAAATCGTTTTCATAGATTTCCGGAGAGGCTATCT 2810
 Qy 2982 GTCAGAAATTTATCTGTTTATCCCGGGTGTAAATCGGAAATTTTGAAGAAATTTAGAAGTGC 3041
 Db 2811 TCCAGAGTTTATCTGTGATTCGCGGTGTAAATGTAGACATTTTCGAAAGTTTAAAGGGCG 2870
 Qy 3042 CATTTACATGCAATCTCCCTATACGATCGGAGAAATGTCGTTAAAAATGCTGATTTTAA 3101
 Db 2871 TATTTTCACTGCAATCTCTCTATATGATGCGAGAAATGTCATTTAAAAACCGTGTATTTCAA 2930
 Qy 3102 TAATGATTTAGCATGCTGGAAATGTAAGAGGCGATGTAGATGT--ACAACAGAGCCATCA 3158
 Db 2931 TAATGGCTTATCATGCTGGAAACGTAAGGCGCATGTAGATTTAGAGAGAACAAACACCA 2990
 Qy 3159 CCGTCTCTGCTTTTATCCAGAAATGGGAAGCAGAAAGTGTCAACAGAGTTCCGCTGTG 3218
 Db 2991 CCGTTCGGTCTCTGTTGTTCCGGAATGGGAAGCAGAAAGTGTCAACAGAGTTTCTGTGTG 3050
 Qy 3219 TCCGGGCGGTGGCTATATCTCTCGTGTACAGCGTACAAAGAGGATATCGAGAGGTTG 3278
 Db 3051 TCCGGGTCGTGGCTATATCTCTGTCAGCGTACAGGAGGGATATGGAGAGGTTG 3110

QY 1425 CATTTCACAACTCTAGGTGTCATGTACACAGTATATCTTTGGACGCCACCGTAGTCAGATCG 1484
 DB 1307 CAGGAAGTGAATCATGGAG--AGCTCCAAATGTTTTCTTTGGACGCCACCGTAGTGAACCCC 1364
 QY 1485 TACAATATACAACTTAGTTTACAGATAGCAATAACAATAACCAATGGTGAATAATCATTCACACCT 1544
 DB 1365 TACAATATACAACTTAGTTTACAGATAGCAATAACAATAACCAATGGTGAATAATCATTCACACCT 1424
 QY 1545 TAATTCAGGTACCTCTCTAGTTCAGTGCAGGCCAGAGATTTTACAGGAGGAGATATAATCCGAAC 1604
 DB 1425 TCAGTCAGGTACTACTGTTTGAAGAGGCCCGGGTTTACGGAGGAGATATCTTTCGACG 1484
 QY 1605 TAACGTTTAATGGTAGTGTACTTAAGTATGGGTCTTAATTTTAAATAATACATCATTTACAGCG 1664
 DB 1485 AACAGTGGAGGACCAATTTGCTTATATACTAATTTAATAATAATAAATGGGCAATACCCCAAG 1544
 QY 1665 GTATCGGTGAGAGTTCGTTTANGTCTGTTCTCAAAACAATGGTCTGAGGAGTAACTGTTCGG 1724
 DB 1545 GTATCGTGCAGAAATACGCTATGCTCTACTACAAATCTAAGAATTTTACGTAACGGTTGC 1604
 QY 1725 AGGAGTACTACTTTTTCATCAAGGATTTCCCTAGTACTATGAGTGCATAATGAGTCTTTGAC 1784
 DB 1605 AGGTGAACGGAATTTTTCGTGCTCAATTTTACAAACAATGATACCGGTACCCCAITTAAC 1664
 QY 1785 ATCTCAATCAATTTAGATTTTGCAGAAATTTCCCTAGTGTATGATGATCTGCGCAGTCAAAC 1844
 DB 1665 ATTCCAATCTTTAGTTACGCAACTAATTAATACAGCTTTTACATTTCCCAATGAGCCAGAG 1724
 QY 1845 TGCT---GGAAATAGTAAGTAATAATGAGGTAGACAAACGTTTCACTTTTGTGATAAAT 1901
 DB 1725 TAGTTTCACAGTAGTGTCTGATCTTTTGTAGTTCAGGGAATGAAGTTTATATATAGACAGATT 1784
 QY 1902 TGAATTCATTTCCAAATTAAGTCACTTCGAGCAGATAGATTTTGAAGGGCGCAAGA 1961
 DB 1785 TGAATTCATTTCCAGTTACTGCAACATTTGAAGCAGATATGATTTGAAGAGAGCACAAA 1844
 QY 1962 GCGGTGAATGCTCTGTTTACTAATACGAATTCGAAGGAATTTGAATAACAGATGTGACAGA 2021
 DB 1845 GCGGTGAATGCTGTTTACTTCTAATAACCAATAGGATATAACAGATGTGACGGA 1904
 QY 2022 TTATCATATTGATCAAGTATCCAAATTTAGTGGCGGTGTTTATCGGATGAATTTGCTTTAGA 2081
 DB 1905 TTATCATATTGATCAAGTATCCAAATTTAGTGGATGTTTATCAGATGAATTTTGTCTGGA 1964
 QY 2082 TGAAGAGAGAGAAATTTCTGGAAGTGAATAATGCGAAACGATCTCAGTGTATGAAGAAA 2141
 DB 1965 TGAAGAGAGAGAAATTTGCGGAAGTGAATAATGCGAAACGATCTCAGTGTATGAGCGGAA 2024
 QY 2142 CTTACTCCAGATCCAAACTTTCATCATCAATTAAGCAACAGATTTCTACTATA 2201
 DB 2025 TTTACTTTCAAGATCCAAACTTTCAGGATCAATAGGCAACTAGAC----- 2070
 QY 2202 TGACCAATCGAATTTTCAATCTATCTCATGAACAACTGTGAACATGGATGGTGGGGAAGTGA 2261
 DB 2071 -----CGTGGTTGGAGGGAAGTAC 2090
 QY 2262 GAACATTTACAATCCAGGAAGGAATGACGTATTTTAAAGAGAAATTTACGTCACACTTACCGGG 2321
 DB 2091 GGATTTACCAATCCAAAGAGGAGATGACGTATTTCAAAGAAAATTTATGTCACTACCGAG 2150
 QY 2322 GACTTTTAAATGAGTGTATCCGAGTATTTTATATCAAAAAATAGGAGATCGGAATTTAAA 2381
 DB 2151 TACCTTTGATGAGTGTATCCAAAGTATTTATATCAAAAAATAGATGAGTCGAAATTTAAA 2210
 QY 2382 AGCTTATCTCGCTACCAATTTAAGAGGGTATATTTGAAGATAGTCAAGATTTAGAGATATA 2441
 DB 2211 ACCCTATCTCGTTTATCAATTTAAGAGGGTATATTCGAGGATAGTCAAGCTTAGAATCTTA 2270
 QY 2442 TTTGATTCGTTTATTAATCGGAACATGAACATTTGATGTTCCAGGTACCGAGTCCGATG 2501
 DB 2271 TTTGATTCGCTTATTAATCGAACAACGAAATGTCGATGAGTACGGGTTCTTTATG 2330
 QY 2502 GCGCTTTTCAGTTGAAAGCCCAATTCGAAGGTGCGGAGAACCGGAATCGATGCGCACCA 2561

DB 2331 GCCGCTTTTCAGTCCAAAGTCCAATCAAGAAAGTGTGGAGAACCGAATCGATGCGGCCACA 2390
 QY 2562 TTTTGAATGAATCCTGATCTAGATTTGTTCTCGACAGATGGAGAAAATGTTGCGCATCA 2621
 DB 2391 CCTTGAATGAATCCTGATCTAGATTTGTTCTCGACAGACGGGAAAATGTTGACATCA 2450
 QY 2622 TTCCCATCTTTCTCTTTGGATATTGATATTGGATGACAGACTTTGCATCAGAACTTAGG 2681
 DB 2451 TTGCGATCATTTCTCTTTGACATTTGATTTGATGTACAGACTTAATCAGGACTTAGA 2510
 QY 2682 CGTGTGGTGGTATTCAAGATTAAAGACGCAAGAGGTCTATGCAAGACTAGGAACTTGGA 2741
 DB 2511 TGTATGGTGTATTTCAAGATTAAAGACGCAAGATGGCCATGCAAGACTAGGAACTTAGA 2570
 QY 2742 ATTTATTGAAGAGAACCACTTATTAGAGAGACACTGCTCTCGTGTGAAGAGCAGAGAA 2801
 DB 2571 GTTTCTCGAAGAGAACCACTTATTGCGGGAGCACTAGCTCGTGTGAAGAGCAGAGAA 2630
 QY 2802 AAAATGGAGAGACAAACGCTGAAAAAATACTAATTTGGAAACAAACAGATATATACAGAGC 2861
 DB 2631 AAAATGGAGAGATAACCTGAAAAATTTGGAATTTGGAACAAATATTGTTTATAAAGAGC 2690
 QY 2862 AAAAGAGCTGTGGATGCTTTTATTGTTGATTTCTCAATATAATAGATTACAAGCGGATAC 2921
 DB 2691 AAAAGAACTGTAGATGCTTTATTGTTTAACTCTCAATATGATCAATTTACAAGCGGATAC 2750
 QY 2922 AAACATTTGGCATCATTCATGCGGCAGATAAATGTTTCTCGAATTCGAGAGGCTTATCT 2981
 DB 2751 GAATATTTGCCATGATTCATGCGGCAGATAAATGTTTCTAGAAATTCGGAAGCGGTATCT 2810
 QY 2982 GTCAGAAATTTATCTGTTATCCCGGTGTAAATGCGGAAATTTTGAAGAAATTAGAAGGTG 3041
 DB 2811 TCCAGAGTTATCTGATTTCCGGGTGTAATGTAGACATTTTGAAGAAATTAAAGGGCG 2870
 QY 3042 CATTTACATGCAATCTCTCTATACGATGCGAAGATGTCGTTTAAAAATGGTGAATTTAA 3101
 DB 2871 TATTTTCACTGCAATCTCTCTATATGATGCGAAGATGTCATTTAAAAACGGTGAATTTCAA 2930
 QY 3102 TAATGATTTAGCATCTGGAATGTAAGGGCATGTAGATGT---ACAACAGAGCCATCA 3158
 DB 2931 TAATGGCTTATCATGCTGGAAACGTAAGGGCATGTAGATGTAGAGAACAAACAAACCA 2990
 QY 3159 CCGTTCCTGCTTTGTTATCCAGAAATGGGAAGCAGAGTGTCAAGCAGTTCGCGCTG 3218
 DB 2991 CCGTTCCGTCCTTTGTTTCCGGAATGGGAAGCAGAGTGTCAAGAAAGTTCGTGTCTG 3050
 QY 3219 TCCGGGGCGTGGCTATATCTCCGTGTCAAGCGGTACAAAGAGGGATATCGAGAGGGTTG 3278
 DB 3051 TCCGGGTGCTGGCTATATCTTCTGTCAAGCGGTACAAAGGGGATATCGGAAGGGTTG 3110
 QY 3279 TGTAAACGATCCATGAAATCGAGAACAAATACAGACGAACATAAATTTAAAACTGTGAAGA 3338
 DB 3111 CGTAACCAATTCATGAGATCGAGAACAAATACAGACGAACATAAATGATTATGCTGCGTAGA 3170
 QY 3339 AGAGGAAGTGTATCCCAACGATACAGAAACGTTAATGATTATATCTGCAACAAGGATAC 3398
 DB 3171 AGAGGAAGTGTATCCCAACCAACACCGTTAATGATTATGCTGCAATCAAGAAAGA 3230
 QY 3399 A-----GCAGTATGTAATTTCCCGTAAATGCTGATGAGGATGATGAGAGTTGATAC 3452
 DB 3231 ATACGGGGTGCCTACACTTTCCCGTAAATCGTGGATATGACGAAACCTTATGGAAGCAATTC 3290
 QY 3453 TACAGCATCTGTTTAAATTTTACAAACCGCTTATGAAGAGAAACGTTATACAGATGTPACGAAG 3512
 DB 3291 TTCTGTACCAAGCTGATTTATGCGTCACTATGAAGAAAATTCGTATACAGATGGAACGAG 3350
 QY 3513 AGATAATCATTTGTAATATACAGAGGGTATGTAATTTATCCCACTACAGCTGGTTA 3572
 DB 3351 AGACAATCTTGTGTAATCTAACAGAGGATATGGGATTTACACCACTACAGCTGGCTA 3410
 QY 3573 TATGACAAAAGAAATTAGAATATCTTCCAGAAACCGATAGGTATGATTTGATTTGAGATTTGAGA 3632

Db 3411 TGTGACAAAGAAATTAGAGTACTTCCAGAAACCGATAAGGTATGGATTGAGATCGGAGA 3470

Qy 3633 AACGGAAGGAAATTTATGAGACGCTGGAATTAATCTCTTATGAGGAA 3684

Db 3471 AACGGAAGGAAATTTATGAGACGCTGGAATTAATCTCTTATGAGGAA 3522

RESULT 10

US-09-873-873-25

; Sequence 25, Application US/09873873

; Patent No. US20020064865A1

; GENERAL INFORMATION:

; APPLICANT: Malvar, Thomas

; APPLICANT: Gilmer, Amy Jelen

; TITLE OF INVENTION: Polynucleotide Compositions Encoding Broad-Spectrum S-Endotoxins

; FILE REFERENCE: MFCO:210--2

; CURRENT APPLICATION NUMBER: US/09/873,873

; CURRENT FILING DATE: 2001-08-20

; PRIOR APPLICATION NUMBER: US 09/253,341

; PRIOR FILING DATE: 1999-02-19

; PRIOR APPLICATION NUMBER: US 08/922,505

; PRIOR FILING DATE: 1997-09-03

; PRIOR APPLICATION NUMBER: US 08/754,490

; PRIOR FILING DATE: 1996-11-20

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 25

; LENGTH: 3534

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Hybrid Delta-Endotoxin

; NAME/KEY: CDS

; LOCATION: (1)...(3531)

US-09-873-873-25

Query Match 38.0%; Score 1402; DB 9; Length 3534;

Best Local Similarity 65.7%; Pred. No. 0;

Matches 2243; Conservative 0; Mismatches 1070; Indels 99; Gaps 10;

Qy 285 TCAGTGGGAAATTTCTAGAACATGTCGAAACATTAATAATCAACAAATAACAGAAA 344

Db 213 TCAATGGAGCGATTTCTGTACAAATGAAACGTTAATNACAAAGAAATAGAAGATT 272

Qy 345 TGCTAGGAATACGGCACTTGCTCGATTACAGGTTTAGGAGATTCTTTTAGAGCCTATCA 404

Db 273 CGCTAGGAACCAAGCCATTTCTAGATTAGAGGACTAAGCAATCTTTATCAAAATTAAGC 332

Qy 405 ACAGTCACTGGAAGATTGGCTAGAAAACCGGTGATGATGCAAGAACGAGAGTGTCTTTA 464

Db 333 AGAATCTTTTAGAGAGTGGGAAGCAGATCCTACTAATCCAGCATTAAGAGAGAGATGCG 392

Qy 465 TACCAATATATAGCCTTAGAACATTGATTTCTTAATCGGATCGGCTTTTCGCAATTAG 524

Db 393 TATTCAATTCATGACATGAACAGTGCCCTTACAAACCGCTATTCTCTTTTGGAGTCA 452

Qy 525 AAACCAAGAAAGTTCATTTAATGATGATGCTCAAGCTGCAAAATTTACACCTATTATT 584

Db 453 AAATTAATCAAGTTCCTTTTATCAGTATATGTTCAAGCTGCAAAATTTACATTTATCAGT 512

Qy 585 ATTGAGAGATGCCCTCTCTTTTGTGATGAAATTTGGGCTTACATCGCAGGAAATTAACG 644

Db 513 TTTGAGAGATGTTTCAAGTGTGAGCAAAAGGTGGGATTTGATGCCGCACTATCAATAG 572

Qy 645 TTATTATGAGCGCAAGTGGGAACAAACGAGAGATTATTCGACTATTTCGCTAGAAATCGTA 704

Db 573 TCGTTAATAAGATTAACTAGGCTTTATTGGCAACTATACAGATCATGCTGTACCGCTGTA 632

Qy 705 TAATACAGGTTCTAAATAGCTTGAGAGGCAAAATGCGCAAGTTGGGTGCGTTATAATCA 764

Db 633 CAATACGGGATTAGAGGCTGTATGGGACCGGATTTCTAGAGATTGGATAGATATAATCA 692

Qy 765 ATTCGCTAGAGATCTAACGTTAGGGGTATTAGATTCTAGTGGCACTATTCCCAAGCTATGA 824

Db 693 ATTTAGAAAGAAATTAACACTAACTGATTATAGATATCGTTCTTATTTCCGAATATGA 752

Qy 825 CACTCGCACATTATCCAAATAAATACGAGTGTCTAGTTAAACAGGGAAGTTTATACAGAGC 884

Db 753 TAGTAGAACGATCCCAATTCGAACAGTTTCCCAATTAACAAGAGAAATTTATACAAACC 812

Qy 885 A--ATTGGAGCAACAGGGGTAATATATGCGAAGTATGAATTTGGTATATATAATATGCACT 942

Db 813 AGTATTAGAAAAATTTTGTATGTTATTTTTCAGGGCTCGGCTCAGGGCATGAAGAAGTAT 872

Qy 943 TCGTTTTTCGCTATAGAGACTGCGGTTATTCGAGGCCGATCTACTTGTATTTTCTAGAA 1002

Db 873 TAGAGTCCCAATTTGATGAGATATCTAACAGTATTAACCATCTATACGGATGCTCATAG 932

Qy 1003 CAACCTTACAAATTTTAGCATTCTATCAGATGGAGTGTCTACTAGGCATATGACTTACTGG 1062

Db 933 GGGTTATTATTATTTGGTCAGGGCATCA-----AATAATGCTCTCTCTGTAGGTT 983

Qy 1063 CGGGGCGACAAATTTCAATCTCGGCCAATAGGAGCGGATTAATATCTCAACGATGGG 1122

Db 984 TTCGGGGCCAGAAATTTCACTTTTCGCTATATGGAACTATGGAA-----ATGCA 1032

Qy 1123 TCTACCAATACTTCTTATTAATCTGTAAAGATTATCACTTCTCTCGAGACGTATATGG 1182

Db 1033 GCTCCACAAACGATTTGTGTCTCACTAGTCAGGGCGGTATAGAACATTTATCGTCC 1092

Qy 1183 ACTGAATCATATGCAAGAGTCTTCTATGGGAAATTTACCTTGAACCTTATTCATGGTGC 1242

Db 1093 ACTTATATAGAACACCTTTTAAATATAGGGATAAATAATCAACAATCTCTGTTCTTGAC 1152

Qy 1243 CCTACTGTTAGATTTAATTTAGAACCTCTCAGAAATCTTTTGAAGAGGTACTGCTAAC 1302

Db 1153 GGGAC-----AGAAATTTGCTTATGGAACTCTCAAAATTTGCCATCCGCTGTATACAGAA 1208

Qy 1303 TATAGTCAACCCCTATGAGTCACTGGGCTTCAATTAAGATTCAGAACTG-AATTACC 1361

Db 1209 AAGCGGAACCGGTAGATTTCGTGATGAATAACCGCCACAGAAATAACACGTGCCACCTAG 1268

Qy 1362 ACCAGAAACAAACAGAACGCAAAATTAATGAATCATATAGTCATAGGTTATCTCATAGG 1421

Db 1269 GCAAGGATTAGTATCATCGATTAGCCATGTTCAATGTTTCGT-----TCAGGCTTTAG 1322

Qy 1422 GCTCATTTCAATCTAGGTGCAATGTAAGTATATCTTTGGACGCAACCGTAGTGAGA 1481

Db 1323 TAAATAGTAGTGAATATAAAGAGCTCCAATGTTTCTTTGGACGCAACCGTAGTGAAC 1382

Qy 1482 TCGTACAAATACCAATAGTTTCAATAGCATTAACAAATACCAATGCTTGAATCAATCAA 1541

Db 1383 CCCTACAAATACCAATGATCCGAGAGGATTAATCAATACCAATGCTTGAATCAATCA 1442

Qy 1542 CCTTAATTCAGGTACCTCTCTAGTTCAGTGGCCAGGATTTACAGGAGGATATAATCCG 1601

Db 1443 ACTTCAGTCACTACTACTGTTAGAGGCGCGGTTTACGGGAGAGATATTTCTTCG 1502

Qy 1602 AACTAACGTTAATGTTAGTGTACTAAGTATGGGTCTTAATTTTAAATATACATCAATACA 1661

Db 1503 ACGAAACAGTGGAGGACCAATTTGCTTATCTATTTGTTAATAATAATGGCAATTAACCCA 1562

Qy 1662 GCGGTATCGGTGAGAGTTCGTTATGCTGTTCTCAACAAATGCTCTGAGGGTAACGT 1721

Db 1563 AAGGTATCGTGAAGAAATACGCTATGCTCTACTACAAATCTAAGAAATTTACGTAACGGT 1622

Qy 1722 CGGAGGAGTACTACTTTTGTATCAAGGATTTCCCTAGTACTATAGTGCATAATGAGTCTTT 1781

Db 1623 TGCAGGTGAACGATTTTTCGTTGCTCAATTTAACAAACATGATACCGGTGACCCATT 1682

Qy 1782 GACATCTCAATCAATTTAGATTTGAGAAATTTCCCTGTAGGTATTAGTGCACTCTGCACTCA 1841

Db 1683 AACATCCCAATCTTTTAGTTACGCAACTAATTAATACAGCTTTTACATTTCCCAATGAGCA 1742

Qy 1842 AACTGCT---GGAAATAGTATAGTAATAATAGCAGGTAGACAAACGTTTCTACTTTGATAA 1898

1743 GAGTAGTTTACAGTAGGTGCTGATCTTTTAGTTTACAGGAATGAAGCTTTATATAGACAG 1802
 1899 AATTGAATTCATTCCTCAATTAAGTCAACTTCGAGCAGAGATACGATTTAGAAAGGGCGCA 1958
 1803 ATTGGAATTCATTCCTCAATTAAGTCAACTTCGAGCAGAGATACGATTTAGAAAGGGCGCA 1862
 1959 AGAGCGGTGAATGCTGTTTAACTAATACGAATCCAAAGAGATTGAAAAACAGATGTGAC 2018
 1863 AAAGCGGTGAATGCTGTTTAACTAATACGAATCCAAAGAGATTGAAAAACAGATGTGAC 1922
 2019 AGATTATCATATGATCAAGATTCATCAATTTAGTGGCGTGTGTTATTCGATGAATTCGCTT 2078
 1923 GGAATTATCATATGATCAAGATTCATCAATTTAGTGGCGTGTGTTATTCGATGAATTCGCTT 1982
 2079 AGATGAAGAGAGAGAAATTTACTTGAAGAGTCAATATGCGAAGAGTCAAGTCAAGTGAAG 2138
 1983 GGAATGAAGAGAGAGAAATTTACTTGAAGAGTCAATATGCGAAGAGTCAAGTCAAGTGAAG 2042
 2139 AAACCTTACTCCAAGATCCAAACTTTCATCATCCATCAATAAGCAACCAAGTTCATATCTAC 2198
 2043 GAATTTACTTCAAGATCCAAACTTTCATCATCCATCAATAAGCAACCAAGTTCATATCTAC 2091
 2199 TAATGAGCAATCGAAATTTTCACTATCTATCTATCTATCTATCTATCTATCTATCTATCT 2258
 2092 -----CGTGGTTGGAGAGGAAG 2108
 2259 TGAGACATTTACATCCAGGAGAGAAATGAGTATTTTAAAGAGAAATTTACCTCAGACTTACC 2318
 2109 TAGCGATTTTACCATCCAAAGAGAGAGATGAGTATTTTAAAGAGAAATTTATGTCACACTACC 2168
 2319 GGGGACCTTTTAAAGAGTGTATTCGAGCTATTTATATCAAAAAATAGGAGAGTCGGAATTT 2378
 2169 AGGTACCTTTTAAAGAGTGTATTCGAGCTATTTATATCAAAAAATAGGAGAGTCGGAATTT 2228
 2379 AAAAGCTTTTACCGCTTATCAATTAAGAGGGTATATCGAAGATAGTCAAGACTTAGAAAT 2438
 2229 AAAAGCTTTTACCGCTTATCAATTAAGAGGGTATATCGAAGATAGTCAAGACTTAGAAAT 2288
 2439 ATATTGATTCGTTTATATCGAAGAGTAAATGAGTATTTTAAAGAGTAAATGAGTAAATGAG 2498
 2289 CTATTTAAATTCGTTTAAATGAGTAAATGAGTATTTTAAAGAGTAAATGAGTAAATGAG 2348
 2499 ATGGCCGCTTTTACGTTTAAAGAGTAAATGAGTATTTTAAAGAGTAAATGAGTAAATGAG 2558
 2349 ATGGCCGCTTTTACGTTTAAAGAGTAAATGAGTATTTTAAAGAGTAAATGAGTAAATGAG 2408
 2559 ACATTTGAATGGAATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 2618
 2409 ACACCTTGAATGGAATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 2468
 2619 TCATTTCCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2678
 2469 TCATTTCCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2528
 2679 AGGCGTGTGGTGTGATTTCAAGATTTAAGACGCGAGAGGTTCATGCAAGATAGGGAATCT 2738
 2529 AGGTGATGGTGTGATTTCAAGATTTAAGACGCGAGAGGTTCATGCAAGATAGGGAATCT 2588
 2739 GGAATTTTGAAGAGAAACCATTTATGAGAGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 2798
 2589 AGAGTTTCTCGAAGAGAAACCATTTATGAGAGAGCACTGCTGCTGCTGCTGCTGCTGCTG 2648
 2799 GAAAAATGAGAGAGAAACCATTTATGAGAGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 2858
 2649 GAAAAATGAGAGAGAAACCATTTATGAGAGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 2708
 2859 GCGAAAGAGAGTCTGATGCTTATTTAGTGTGATTTCTCAATATAATAGATTACAAGCGGA 2918
 2709 GCGAAAGAGAGTCTGATGCTTATTTAGTGTGATTTCTCAATATAATAGATTACAAGCGGA 2768
 2919 TACAAACATTTGGCATGATTTATCGGGGAGATTAACCTGTTTCATCGAATTCGAGAGGCTTA 2978
 2769 TACAAACATTTGGCATGATTTATCGGGGAGATTAACCTGTTTCATCGAATTCGAGAGGCTTA 2828

2979 TCTGTGAGAAATATCTGTTATCCGGGTGTAAATCGGAAATTTTGAAGAAATTAGAAAGG 3038
 2829 TCTGCTGAGCTGTCTGTGATTCGGGTGTCAATCGGCTATTTTGAAGAAATTAGAAAGG 2888
 3039 TCGCAATTTACATGCAATCTCCCTATATACGATCGAGAAATGTCTTTAAAAATGGTGAATTT 3098
 2889 GCGTATTTTACATGCAATCTCCCTATATGATGCGAGAAATGTCTTTAAAAATGGTGAATTT 2948
 3099 TAATAATGATTTAGCATGCTGGAATGTAAAGGGCATGTAGATGT--ACAACAGAGCCCA 3155
 2949 TAATAATGCGTTATCTGCTGGAACGTGAAGGGCATGTAGATGTGAAGAACCAACAA 3008
 3156 TCACCGCTTCTGCTTCTGTTATCCAGAAATGGGAAGCAGAGTGTCAACAGCAGTTCGCGT 3215
 3009 CCAACGTTGCTGCTTCTGTTTCCGGAATGGGAAGCAGAGTGTCAACAGCAGTTCGCGT 3068
 3216 CTGTCCGGGGCGTGGCTATATCTCCGTGTCAACAGCTTCAAAAGAGGGATATGAGAGGG 3275
 3069 CTGTCCGGGGCGTGGCTATATCTCCGTGTCAACAGCTTCAAAAGAGGGATATGAGAGGG 3128
 3276 TTGTGTAAACGATCCATGAATCGAGAAACATACAGAGCAACTAAAAATTTAAACCTGTGA 3335
 3129 TTGCGTAAACGATTCATGAGATCGAGAAACATACAGAGCAACTAAAAATTTAGCAACTGCGT 3188
 3336 AGAAGAGGAGTGTATCCCAACGGATACAGAAACGTGTAAATGATTATCTGCACACCAAGG 3395
 3189 AGAAGAGGAGTGTATCCCAACGGATACAGAAACGTGTAAATGATTATCTGCATTAATCAAGA 3248
 3396 TACAGCAGTATGTAATTTCCGCTTAATCTGATATGAGGATGCATATGAAAGTGTGATCTAC 3455
 3249 AGAATACGGAGGTGCGT-----ACACTTCTCTAATCGAGGATATAACGAAGCTCTTCT 3302
 3456 AGCATCTGTTTAAATTAACAAACCGACTTATGAAGAGAAACGTATACAGATGTACGAGAGA 3515
 3303 CGTACCAGCTGATTTATGCGTCTATGAAGAAAAATCGTATACAGATGTACGAGAGA 3362
 3516 TAATCAATTTGAATATGACAGAGGGTGTGTGAATTTATCAACCACTACCACTGCTGTTATAT 3575
 3363 GAATCTTGTGATTTAAACAGGGGTATAGGGATTAACGCCCACTACCACTGCTGTTATAT 3422
 3576 GACAAAGAAATTAAGATACTTCCAGAAACCGGATAGGATGGAATGAGATTCGAGATTCGAGAAC 3635
 3423 GACAAAGAAATTAAGATACTTCCAGAAACCGGATAGGATGGAATGAGATTCGAGATTCGAGAAC 3482
 3636 GGAAGGAGTATTTATGATAGCAGCGTGAATTTACTCTTATGAGAGGAATAG 3687
 3483 GGAAGGAGTATTTATGATAGCAGCGTGAATTTACTCTTATGAGAGGAATAG 3534

RESULT 11

US-09-916-956A-25
 ; Sequence 25, Application US/09916956A
 ; Publication No. US2003001751A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Malvar, Thomas
 ; APPLICANT: Gilmer, Amy Jelen
 ; TITLE OF INVENTION: Polynucleotide Compositions Encoding Broad-Spectrum S-Endotoxins
 ; FILE REFERENCE: MECO:211--1
 ; CURRENT APPLICATION NUMBER: US/09/916,956A
 ; CURRENT FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: US 09/253,331
 ; PRIOR FILING DATE: 1999-02-19
 ; PRIOR APPLICATION NUMBER: US 08/922,505
 ; PRIOR FILING DATE: 1997-09-03
 ; PRIOR APPLICATION NUMBER: US 08/754,490
 ; PRIOR FILING DATE: 1996-11-20
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 25
 ; LENGTH: 3534
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence

FEATURE:
; OTHER INFORMATION: Hybrid Delta-Endotoxin
; NAME/KEY: CDS
; LOCATION: (1) .. (3531)
US-09-916-956A-25

Query Match	38.0%;	Score 1402;	DB 10;	Length 3534;
Best Local Similarity	65.7%;	Pred. No. 0;	Mismatches 1070;	Indels 99;
Matches 2243;	Conservative 0;			Gaps 10;

Qy	285	TCAGTGGGAAATTTCTCTAGAACATGTCGAAACATTTATAATCAACAAATAACAGAAAA	344
Db	213	TCATGGGACCATTTCTGTAGCAATTTGAACAGTTAATTAACCAAGAAATGAGAAAT	272
Qy	345	TGCTAGGAATACGGACATGCTCGATTAACAGTTTAGGAGATTCCTTTAGAGCCTATCA	404
Db	273	CGCTAGGAACCAAGCCATTTCTAGATTAGAGGACTAAGCAATCTTTATCAAAATTTAGCG	332
Qy	405	ACAGTCACTTGAAGATTGGCTAGAAACCGTGATGATGCAAGAACGAGAGTGTCTTTTA	464
Db	333	AGAATCTTTTAGAGAGTGGGAGGAGATCTCTACTAATCCAGCATTAAAGAGAGATGCG	392
Qy	465	TACCAATATATAGCCTTTAGAACATTTGATTTCTTAATGCGATGCGCTTTTCGCAATTAG	524
Db	393	TATTCATTTCAATGACATGAACAGTGCCCTTACACCGCTATTCCTCTTTTTCGAGTTCA	452
Qy	525	AAACCAAGAGTTCCATTTATTAATGGTATATGCTCAAGCTGCAAAATTTACACCTATTAT	584
Db	453	AAATATCAAGTTCCTCTTTTATCAGTATATGTTCAAGCTGCAAAATTTACATTTATCAGT	512
Qy	585	ATTGAGAGATCCCTCTCTTTTGGTAGTGAATTTTGGGCTTACATCGCAGGAAATTTCAAG	644
Db	513	TTTGAGAGATGTTTCAGTGTTTGGCAAGAGTGGGGATTTGATGCCGCGACTATCAATAG	572
Qy	645	TTATTATGAGCCCAAGTGGAAACAAACGAGAGATTTTCCGACTATTTCGCTAGATGTA	704
Db	573	TCGTTTAAATGATTTAACTAGGCTTTATGGCACTATACAGATCATGCTGTACGCTGGTA	632
Qy	705	TAATACAGGTCTAAATAGCTTTGAGAGGACAAATGCGCAAGTTGGGTGCGTTATAATCA	764
Db	633	CAATACGGGATTAGAGCGTGTATGGGACCGGATTTCTAGAGATTGGATAAGATATAATCA	692
Qy	765	ATTCGGTAGAGATCTAAGTTAGGGGTATTAGATCTAGTGCGCACTATTCCCAAGCTAGA	824
Db	693	ATTTAGAGAGAAATTAACACTAATCTGTATTAGATATCGTTCTCTATTTCGCAACTATGA	752
Qy	825	CACCTGCATTTATCCATAAATACGAGTCTCAGTTAACAGGGAAGTTTATACAGAGC	884
Db	753	TAGTAGAACGTATCCAAATTCGAACAGTTTCCCAATTAACAGAGAAATTTTACAAACCC	812
Qy	885	A--ATTGGAGCAACAGGGGTAAATATATGCAAGTATGAATTTGGTATAATAATATGCACT	942
Db	813	AGTATTAGAAATTTTGATGTAGTTTTCGAGGCTCGGCTCAGGGCATAGAAAGATAT	872
Qy	943	TCGTTTTCGGCTATAGAGACTCGGGTTATCCGAAGCCCGCATCTACTTGAATTTTCTAGAA	1002
Db	873	TAGGAGTCCACATTTGATGGATATCTTAAACAGTATACCATCTATACGGATGCTCATAG	932
Qy	1003	CAACTTACAATTTTATGACATTCATCAGATGGAGTGTCTACTAGGCATATGACTTCTGG	1062
Db	933	GGGTATTATTATTGTTGGTCAGGGCATCA-----AATATGGCTTCTCTGTAGGGTT	983
Qy	1063	CGGGGGCACAAATTCATCTCGGCCAATAGGAGCGGATTAATACCTCAACCGATGGG	1122
Db	984	TTCCGGGCGCAATTTCACTTTTCGCTATATGGAACTATGGGAA-----ATGCA	1032
Qy	1123	TCTACCAATATCTCTATTAAATCTGTAGATTTATCTTCTCTCGAGAGCTATATTGG	1182
Db	1033	GCTCCAGAACCAACGATTTGTTGCTCAACTAGGTTCAGGCGGTATAGAACATTTATCGTCC	1092
Qy	1183	ACTGAATCATATGAGGAGTCTCTATATGGGGAATTTACCTTGAACCTATTATGTTGTC	1242
Db	1093	ACTTTATATAGAGACCTTTTAAATATAGGGATAAATAATCAACACTATCTGTTCTTGAC	1152

Qy	1243	CCTACTGTTAGATTTAATTTTAGGAACCCCTCAGAAATACTTTTTGAAAGAGGTACTGCTAAC	1302
Db	1153	GGGAC-----AGAAATTTGCTTATGGAACCTCTCAAAATTTGCCATCTGTATACAGAAA	1208
Qy	1303	TATAGTCAACCCCTATGAGTCACCTGGGCTTCAATTTAAAGATTTCAGAACTG-AATTACC	1361
Db	1209	AAGCGGAACGGTAGATTTCGCTGGATGAATACCCGCCAGAAATACACAGTGGCCACCTAG	1268
Qy	1362	ACCAGAAACACAGAACGACCAAAATTTAGATTCATATAGTCTATAGGTTTATCTCATAGG	1421
Db	1269	GCAAGGATTTAGTCATCGATTAAGCCATGTTCAATGTTTCT-----TCAGGCTTTAG	1322
Qy	1422	GCTCATTTCAAACTTAGGGTGCATGACAGTATATTTCTTGGACGACCGCTAGTGCAGA	1481
Db	1323	TAATAGTAGTGTAGTATAATAGAGCTCCAATGTTTCTTGGAGCGCACCGCTAGTGCAC	1382
Qy	1482	TCGTACAAATACCAATTTAGTTCAGATAGCATACAAATACCAATTCGTTAAATCATTTCAA	1541
Db	1383	CCCTACAAATACAAATTCATCCGAGAGGATTTACTCAAAATACCATTTGGTAAAGCACATAC	1442
Qy	1542	CCTTAATTCAGGTACCTCTGTAGTCAGTGGGCCAGGATTTTACAGAGGGGATATAATCCG	1601
Db	1443	ACTTCAGTCAGTACTACTGTTTGAAGGGCCCGGTTTACGGGAGAGATATTTCTCG	1502
Qy	1602	AACCTAACGTTTAATGGTAGTGTACTAAGTATGGGTCTTTAAATTTTAAATACATCATTA	1661
Db	1503	ACGAAACAGTGGAGGACCAATTTGCTTATCTATTTGTTAAATATAAATGGCAATTAACCCA	1562
Qy	1662	CGGCTATCGGTGAGAGTTCGTTATGCTGCTTCTCAAAACATGGTCTCGAGGGTAACGT	1721
Db	1563	AAGGTATTCGTCAGGAATACGCTATGCTCTACTACAAATCTAAGAAATTTTACGTAACCGT	1622
Qy	1722	CGAGGGAGTACTACTTTTGATCAAGGATTCCTTAGTACTATGAGTGCAAATGAGTCTTT	1781
Db	1623	TGCAAGTGAACGGATTTTGTGCTCAATTTACAAACAAATGATACCGGTGACCCATTT	1682
Qy	1782	GACATCTCAATCAATTTAGATTTGCAGAAATTTCTGTAGTATTAGTGCATCTGGCAGTCA	1841
Db	1683	AACATTCCAATCTTTTAGTTACGCAACTATTAAACAGCTTTTACATTTCCCAATGAGCCA	1742
Qy	1842	AACGTCT---GGAATAAGTATAGTATAATGAGGTAGACAAACGTTTCACTTTGATAA	1898
Db	1743	GAGTAGTTTACAGTAGGTGCTGATCTTTTAGTTCAGGGAAATGAAGTTTATATAGACAG	1802
Qy	1899	AATTTGAATTCATTCCAATTAACGACCTTCGAGAGCAGATACGATTTAGAAAGGGCGCA	1958
Db	1803	ATTTGAATTTGATTCAGTTACTGCAACATTTGAAGCAGAAATGATTTTAGAAGAGCACA	1862
Qy	1959	AGAGCGGTGAATGCTCTGTTTACTAATACGAATCCAAAGATTCGAAACAGATGTGAC	2018
Db	1863	AAAGCGGTGAATGCGCTGTTTACTTCTATAAACCAATAGGGATATAAACAGATGTGAC	1922
Qy	2019	AGATTATCATTTGATCAAGTATCCAAATTTAGTGGCGTGTATTCGGATGAATTTCTGCTT	2078
Db	1923	GGATTATCATTTGATCAAGTATCCAAATTTAGTGGATTTTATCAGATGAATTTTGTCT	1982
Qy	2079	AGATGAAGAGAGAAATTTCTTGAAGAGTCAATATGCGAAACGACTCAGTGTATGAAG	2138
Db	1983	GGATGAAGAGCGGAAATTTGTCGGAAGAGTCAAACTTCGGAAGCGACTCAGTGTATGAGC	2042
Qy	2139	AAACTTACTCCAGATCCAACTTCACATCCATCAATAAGCAACAGACTTTCATATCTAC	2198
Db	2043	GAATTTACTTTCAGATCCAACTTCGAAGGATCAATAGGCAACTAGAC-----	2091
Qy	2199	TAATGAGCAATCGAAATTTCACTCTATCCATGAACAACTCTGAACATCGATGTGGGGAAG	2258
Db	2092	-----CGTGGTGGAGAGGAAG	2108
Qy	2259	TGAGAACTTACATCCAGGAAGGAATGAGTATTTTAAAGAGATTTACCTCACACTACC	2318
Db	2109	TACGGATATTACCATCCAAAGAGGAGATGAGTATTCAAAGAAATTTATGTCACTACC	2168

QY 2319 GGGACCTTTTAATGAGTGTATCCGAGTATTTATATCAAAAAATAGGAGAGTCGGAATT 2378
 Db |||||
 QY 2169 AGGTACCTTTTGATGAGTGTATCCAAATATTTGATCAAAAATCGATGAATCAAAATT 2228
 Db |||||
 QY 2379 AAAAGCTTTATCTGCTACCAATTAAGAGGGTATATTGAAGATAGTCAAGATTTAGAGAT 2438
 Db |||||
 QY 2229 AAAAGCTTTATCCGCTTATCAATTAAGAGGGTATATCGAAGATAGTCAAGATTTAGAAAT 2288
 Db |||||
 QY 2439 ATATTGATTCGTTTATTAATCGAAACATGAACATTCGATGTTCCAGTACCGAGTCGCT 2498
 Db |||||
 QY 2289 CTATTTATTCGCTCAATCGAAACATGAACATTAATGTCCAGGTACGGGTTCCTT 2348
 Db |||||
 QY 2499 ATGCCCGCTTTACGTTGAAAGCCCAATCGGAAGTGCAGAAACCGAATCGATCGCAGC 2558
 Db |||||
 QY 2349 ATGCCCGCTTTACGCCCAAGTCCAAATCGGAAGTGTGGAGAGCGGAATCGATCGCGGCC 2408
 Db |||||
 QY 2559 ACATTTGATGGAATCCTGATCTAGATTTGTTCTTCGAGAGATGGAGAAAAATGTGGCGA 2618
 Db |||||
 QY 2409 ACACCTTGAATGGAATCCTGATCTAGATTTGTTCTGTAGGATGGAGAAAAATGTGGCCA 2468
 Db |||||
 QY 2619 TCATTCCTCATCTTTCTTTGGATATTTGATTTGGAATGCACAGACTTGCATGAGAACT 2678
 Db |||||
 QY 2469 TCATTCGATCATCTTTCTTTAGACATTTGATGAGATGTACAGACTTAATGAGGACCT 2528
 Db |||||
 QY 2679 AGCGTGTGGTGTATTTCAAGATTAAGACGCGAGGAAGGTCAATCAAGACTTAGGGAATCT 2738
 Db |||||
 QY 2529 AGGTGTATGGTGTATTTAAGATTAAGACGCGAGATGGCAAGACTTAGGGAATCT 2588
 Db |||||
 QY 2739 GGAATTTATGAAGAGAAACCATTTATGAGAGAGCACTGTCTGTGTGAAGAGAGAGA 2798
 Db |||||
 QY 2589 AGATTTCTCGAAGAGAAACCATTTAGTAGGAGAGCGCTAGCTGTGTGAAAAGAGCGGA 2648
 Db |||||
 QY 2799 GAAAAATGAGAGACAAACGTGAAACATCAATTTGGAACAAACAGATATATACAGA 2858
 Db |||||
 QY 2649 GAAAAATGAGAGACAAACGTGAAACATTTGGAATGGAACAAATATCGTTTATAAGA 2708
 Db |||||
 QY 2859 GCGAAAAAGAGCTGTGGATGCTTTATTTGTAGATTTCTCAATATATAGATTTACAAGCGGA 2918
 Db |||||
 QY 2709 GCGAAAAAGATCTGTAGATGCTTTATTTGTAACTCTCAATATGATCAATTTACAAGCGGA 2768
 Db |||||
 QY 2919 TACAAACATTTGGCATGATTTCTGCGGAGATATAAATCTGTTTCATCGAATTCGAGGGTTA 2978
 Db |||||
 QY 2769 TACGAATTTGCCATGATTTCTGCGGAGATATAAATCGTTTCATAGCATTCGAGAGCTTA 2828
 Db |||||
 QY 2979 TCTGTCAGAATTTATCTGTATCTCCGCTGTAAATGCGGAAATTTTCAAGATTTAGAAAG 3038
 Db |||||
 QY 2829 TCTGCTGAGCTGTCTGTGATTTCCGCTGTCAATGCGGCTATTTTGAAGATTTAGAAAG 2888
 Db |||||
 QY 3039 TCGCATTTATCACTGCAATCTCCCTATACGATGCGAGAAATGTCTGTTAAAAATGTGATTT 3098
 Db |||||
 QY 2889 GCGTATTTTCACTGCAATCTCCCTATATGATGCGAGAAATGTCAATTTAAAAATGTGATTT 2948
 Db |||||
 QY 3099 TAAATATGGAATAGCATGCTGGATGTAAAGGGCATGTAGATGT---ACAACAGAGCCA 3155
 Db |||||
 QY 2949 TAAATATGSGCTTATCTGCTGGAAACGTGAAAGGGCATGTAGATGTAGAGAACAAACAA 3008
 Db |||||
 QY 3156 TCACCGTTCCTGCTGTTATCCAGAAATGCGAAGAGAGAGTGTCAACAGCATTTCCGCT 3215
 Db |||||
 QY 3009 CCAACGTTTCGCTGCTGTTGTTCCGGAATGGGAAGAGAGTGTCAACAGAGTTCTGT 3068
 Db |||||
 QY 3216 CTGTCGGGGCGTGGCTTATATCTCCGTGTACAGCGTACAAAGAGGGATATGGAGAGG 3275
 Db |||||
 QY 3069 CTGTCGGGGCGTGGCTTATATCTTCGTGTACACAGCGTACAAAGAGGGATATGGAGAGG 3128
 Db |||||
 QY 3276 TTGTGTAACCATTCATGAATTCGAGACAAATACAGAGCACTAAATTTTAAACCTGTA 3335
 Db |||||
 QY 3129 TTGCGTAACCATTCATGAGATTCGAGAACAAATACAGAGCACTGAAGTTTATGAACTGCGT 3188
 Db |||||
 QY 3336 AGAAGAGGAAGTGTATCCACCGGATACAGAAACGTGTAATGATTTATCTGCAACCAAGG 3395
 Db |||||
 QY 3189 AGAAGAGGAATCTATCCAAATACACGGTAACTGTAAATGATTTACTGTAAATCAAGA 3248
 Db |||||
 QY 3396 TACAGCATATGTAATTTCCCGTAATGCTGGATATGAGGATGCATATGAAGTTGATACTAC 3455
 Db |||||

Db 3249 AGAATACGGAGGTGCGT-----ACATTCCTCGTAATCGAGGATATAACGAAGCTCTTC 3302
 QY 3456 AGCATCTGTTAAATTAACAACCGACTTATGAAGAAACGATATACAGATGTACGAAGAGA 3515
 Db 3303 CGTACACAGCTGATGCTGCTATGAAAGAAAAATCGTATACAGATGGAAGAGAGA 3362
 QY 3516 TAATCATTTGAAATATACAGAGAGGTATGTAATTTATCCACCACTACCACTGGTTATAT 3575
 Db 3363 GAATCTTGTGAATTTAAACAGAGGGTATAGGATTTACACGCCACTACCACTGGTTATGT 3422
 QY 3576 GACAAAAGAAATTAAGATTAATTCACAGAAACCGAATAGGATTTGAGATTTGAGAAAC 3635
 Db 3423 GACAAAAGAAATTAAGATTAATTCACAGAAACCGAATAGGATTTGAGATTTGAGAAAC 3482
 QY 3636 GGAAGGGAAGTTTATTTGTAGACAGCGTGGAAATTTACTCTTTATCGGGAATAG 3687
 Db 3483 GGAAGGGAACATTTATCTGTGACACGCGTGGAAATTTACTCTTTATCGGGAATAG 3534

RESULT 12
 US-09-997-914-25
 ; Sequence 25, Application US/09997914
 ; Publication No. US20030119158A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Malvar, Thomas
 ; APPLICANT: Gilmer, Amy Jelen
 ; TITLE OF INVENTION: Polynucleotide Compositions Encoding Broad Spectrum d-Endotoxins
 ; FILE REFERENCE: 11792.0215.DVUS01 MECO:215--1
 ; CURRENT APPLICATION NUMBER: US/09/997,914
 ; CURRENT FILING DATE: 2001-11-30
 ; PRIOR APPLICATION NUMBER: US 09/261,040
 ; PRIOR FILING DATE: 1999-03-02
 ; PRIOR APPLICATION NUMBER: US 08/754,490
 ; PRIOR FILING DATE: 1996-11-20
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 25
 ; LENGTH: 3534
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Hybrid Delta-Endotoxin
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(3531)
 ; US-09-997-914-25

Query Match 38.0%; Score 1402; DB 10; Length 3534;
 Best Local Similarity 65.7%; Pred. No. 0;
 Matches 2243; Conservative 0; Mismatches 1070; Indels 99; Gaps 10;

QY 285 TCAGTGGGAATTTTCTAGAACATGTCGAACAACTTATAAATCAACAAATAACAGAAA 344
 Db 213 TCATGGAGCGCAATTTCTGTACAAATGAACTGTAATTAACCAAGAATAGAGAAAT 272
 QY 345 TGCTAGGAATACGCGCACTTGTCTGATTACAGGTTTAGGAGATTTCTTTAGAGCCTATCA 404
 Db 273 CGCTAGGAACCAAGCCATTTCTAGATTAGAGGACTAAGCAATCTTTATCAAAATTTACGC 332
 QY 405 ACAGTCACTTGAAGATTTGGCTAGAAAACCGTGTATGATGCAAGAACGAGAGATGTTCTTTA 464
 Db 333 AGAATCTTTTAGAGAGTGGGAACGAGATCCTACTAATTAATCCAGCATTAAGAGAGAGATGCG 392
 QY 465 TACCAATATATAGCTTGAACCTTGAATTTCTTAATGCGATCCGCTTTTCGCAATTTAG 524
 Db 393 TATTCAATTCATGACATGAAACAGTCCCTTACACCGCTATTTCTCTTTTTCGAGTTCA 452
 QY 525 AAACCAAGAGTTTCCATTTATTAATGTTATATGCTCAAGCTGCAAAATTTACACCTATTATT 584
 Db 453 AAATTTCAAGTTCTCTTTTATCAGTATATGTTCAAGCTGCAAAATTTACATTTATCAGT 512
 QY 585 ATTGAGAGATGCTCTCTTTTGTGTTAGTGAATTTGGCTTACATCGCAGAAATTCACG 644
 Db |||||

Db 513 TTTGAGAGATGTTTTCAGTGTGTTGGACAAAGGTGGGATTTGATCGCGGACTATCAATAG 572
Qy 645 TTAATTATGAGCCGCAAGTGGAAACAAACGAGAGATTAATTCCGACTAATTGCGTGTAGATGGTA 704
Db 573 TCGTTAATAATGATTAATACTAGGCTTAATTGGCAACTATACAGATCATGCTGTACGCTGGTA 632
Qy 705 TAATACAGGTCTTAATAGCTTGAGAGGGACAAAGTCCGCAAGTTGGGTGGCTTAATATCA 764
Db 633 CAATACGGGATTAGAGCGGTGATGGGACCGGATCTTAGAGATTGGATATAGATATATCA 692
Qy 765 ATTCGGTAGAGATCTAACGTTTAGGGTATTAGATCTAGTGCCACTAATCCCAAGCTATGA 824
Db 693 ATTTAGAAGAGAAATTAACACTAATCTGTATTAGATATCGTTCTCTATTTCGCACTATGA 752
Qy 825 CACTCGCACTTAATCAATAAATACGAGTCTCAGTTAACAGGGAAGTTTATACAGACGC 884
Db 753 TAGTAGAACGTTATCCAAATTCGAAACAGTTTCCCAATTAACAGAGAAATTTATACAAACCC 812
Qy 885 A--ATTGGAGCAACAGGGTAAATATGCGCAAGTATGAATTGGTATATATATATATGCACT 942
Db 813 AGTATTAGAAAATTTGATGGTAGTTTTCGAGGCTCGGCTCAGGGCATAGAAAGATAT 872
Qy 943 TCGTTTTCCGCTATAGAGACTGCGGTTATCCGAAGCCGCACTACTCTGATTGTTTCTAGAA 1002
Db 873 TAGGAGTCCACATTTGATGGATATACTTAACAGTATAACCACTATACGATGCTCATAG 932
Qy 1003 CAACCTTACAAATTTTACACTTTCATCAGATGAGTGTCTACTAGGCATATGACTTACTGG 1062
Db 933 GGGTTAATTATTGGTCAGGGCATCA-----AATAATGGCTTCTCCTGTAGGGTT 983
Qy 1063 CGGGGGCACAAATCAATCTCGGCCAATAGGAGCGGATTAATATCTCAACGCAATGGG 1122
Db 984 TTCGGGGCCAGAAATTCATTTTCGCTATATAGGAACATATGGGA-----ATGCA 1032
Qy 1123 TCTACCAATATCTTATTAATCTGTGAAGATTATCATTTCTCTCTCGAGAGGTATATGG 1182
Db 1033 GCTCCACAAACGTAATTTGTTGCTCACTAGGTGAGGGGTGTATAGAACATTAATCTGTC 1092
Qy 1183 ACTGAATCATATGACGAGTGTCTTCTATGGGAATTTACCTTGAACCTATTTCATGTTGTC 1242
Db 1093 ACTTTATATAGAACCTTTTAAATATAGGGATAAATAATCAACAACTATCTGTTCTTGAC 1152
Qy 1243 CCTACTGTTAGATTAATTTTAGAACCTCAGAAATACTTTTGAAGAGGTACTGTCTAAC 1302
Db 1153 GGGAC-----AGAATTTGCTTATGGAACCTCTCAAAATTTGCCATCCGCTGTATACAGAA 1208
Qy 1303 TATAGTCAACCTATGATGACCTGGGCTTCAATTAAGAATTCAGAAACCTG-AAATTACC 1361
Db 1209 AAGCGGAACGGTAGATTGCTGGATGAATATACGCCACAGAAATCAACAGTGCACCTAG 1268
Qy 1362 ACCAGAAACAAACAGAACGCAAAATTAATGAATCATATAGTCATAGGTTATCTCAATAGG 1421
Db 1269 GCAAGGATTTAGTCAATCGATTAAAGCAATGTTCAATGTTTCGT-----TCAGGCTTTAG 1322
Qy 1422 GCTCATTTCAAACTAGGTGCAATGATACGATATATTTTGGACGACCGTAGTGCGAGA 1481
Db 1323 TAATAGTAGTGTAAAGTAAATAGAGCTCCAATGTTTTTCTTGGACGACCGTAGTGAAC 1382
Qy 1482 TCGTACAAATACCAATAGTTTCAAGATAGCATACAAATACCAATGCTTGAATCAATTCAA 1541
Db 1383 CCTTACAAATACAAATGATCGGAGAGATTAATCAATACCAATGCTTGAATGAAGACATAC 1442
Qy 1542 CCTTAAATTCAGGTACCTCTGTAGTCACTGGGCCAGGATTTACAGAGGGGATATAATCCG 1601
Db 1443 ACTTCAGTCAGTACTACTGTTGAAGGGCCCGGTTTACGGGAGAGATATCTTCG 1502
Qy 1602 AACTAAAGTTAATCGGTAGTGTACTAAGTATGGGTCTTAATTTTAATATACATCAATTACA 1661
Db 1503 ACAGAACAGTGGAGGACCAATTTGCTTATCTACTATTGTTAATATAAATGGGCAATTACC 1562
Qy 1662 CGGCTATCGGTAGAGTTCGTTATGCTGCTTCTCAAAACAAATGCTCTCAGGGTAACTGT 1721
Db 1563 AAGGTATCGTCAAGAAATACGCTATGCTCTACTTACAAATCTAGAAATTTACGTAACGGT 1622

Qy 1722 CGGAGGAGTACTACTTTTGTATCAAGGATTCCTTAGTACTATAGTGTCAAAATGATCTTT 1781
Db 1623 TGCAGGTGAACGGGATTTTGTGCTGTCATTTAAACAAACAATGATACCGGTGACCCCAT 1682
Qy 1782 GACATCTCAATCAATTTAGATTTGCAGAAATTTCTGTAGGTATTAGTCATCTGCGAGTCA 1841
Db 1683 AACATTCCAATCTTTTGTAGTACGCAACTATTAATACAGCTTTTATCATTTCCCAATGAGCCA 1742
Qy 1842 AACTGCT---GGAATAAGTATAAGTATAATGCAAGGTAGACAAACGTTTTCATTTGTATAA 1898
Db 1743 GAGTAGTTTTCACAGTAGTGTCTGATACTTTTGTAGTTAGGGAATGAAGTTTATATAGACAG 1802
Qy 1899 AATTGAATTCATTCCAATTAATCTGCAACCTTCGAAAGCAAGATACGATTTAGAAAGGGCGCA 1958
Db 1803 AATTGAATTTGATTCAGTTACTGCAACATTTGAAGCAGAAATATGATTTAGAAAGAGCACA 1862
Qy 1959 AGAGCGGTCAATGCTCTGTTTACTTAATACGAATCCAAAGATTTGAAACAGATGTGAC 2018
Db 1863 AAGGCGGTGAATGCGCTGTTTACTTCTATTAACCAATAGGGATTAACACAGATGTGAC 1922
Qy 2019 AGATTATCATATTGATCAAGTATTCCAATTTAGTGGCGTGTTTATCGGATGAATTTCTGCT 2078
Db 1923 GGAATATCATATTGATCAAGTATCCAAATTTAGTGATTTGTTATCAGATGAATTTGTCT 1982
Qy 2079 AGATGAAAAGAGAGAAATTAATCTGGAAGGTGAATATTCGAAACGACTCAGTGTATGAAG 2138
Db 1983 GGAATGAAAAGAGAGAAATTTGTCGGAAGGTCAAAACATCGGAAGCGACTCAGTGTATGACG 2042
Qy 2139 AACTTACTTCCAGATCCAACTTCATCCATCAATAAGCAACCAAGCTTCATATCTAC 2198
Db 2043 GAATTTACTTCAAGATCCAACTTCAAGGATCAATAGGCAACTAGAC----- 2091
Qy 2199 TAATGAGCAATCGAAATTTACATCTATCCATGAACAACTCTGAACATCGATGGTGGGGAAG 2258
Db 2092 -----CGTGGTTGGAGAGAGAG 2108
Qy 2259 TGAGAACATTTCAATCCAGGAAGAAATGACGTATTTTAAAGAGAAATTAACGTCACTACC 2318
Db 2109 TACGATATTTACCATCCAAAGAGAGATGACGTATTCAGAAAGAAATTTATGTCACTACC 2168
Qy 2319 GGGGACTTTTAATGAGTGTATCCGAGTATTTATATCAAAATATAGAGATCGGAAT 2378
Db 2169 AGGTACCTTTGATGAGTGTATCCAAATATTTGTATCAAAATATCGATGAATCAAAAT 2228
Qy 2379 AAAAGCTTATACCTGCTACCAATTAAGAGGTATATTGAAGATAGTCAAGATTTAGAGAT 2438
Db 2229 AAAAGCTTTTACCGCTTATCAATTAAGAGGTATATCGAAGATAGTCAAGATTTAGAAAT 2288
Qy 2439 ATATTTGATTCGTTATTAATCGAAACATGAACATTTGGATTTCCAGGTACCGAGTCCGT 2498
Db 2289 CTATTTAATTCGCTACAATGCAAAACATGAACACAGTAAATGTGCCAGGTACGGGTTCCT 2348
Qy 2499 ATGGCGCTTTTCAAGTGAAGCCCAATCGGAAGTGGGAGAACCGAATCGATCGGCACC 2558
Db 2349 ATGGCGCTTTTACGCCCAAGTCCAATCGGAAGTGGGAGAGCCGAATCGATCGGCACC 2408
Qy 2559 ACATTTGAAATGAAATCCTGATCTAGATTTCTCTGACAGAGATGGAGAAAATCTGGGCA 2618
Db 2409 ACACCTTGAAATGAAATCCTGATTTAGATTTCTGTTAGGATGGAGAAAATCTGGGCA 2468
Qy 2619 TCATTTCCCATTTCTTTTGGATATTTGATTTGATGCAAGCTTGCATGAGAAATCT 2678
Db 2469 TCATTTCCCATTTCTTTTGGATATTTGATTTGATGCAAGCTTGCATGAGAAATCT 2528
Qy 2679 AGGCGTGGGTGATTTCAAGATTAAGACGAGGAAGGTTCATGCAAGCTAGGGAATCT 2738
Db 2529 AGGTGATGGGTGATCTTTAAGATTAAGACGCAAGATGGGCAAGCTAGGGAATCT 2598
Qy 2739 GGAATTTATTAAGAGAAACCAATTTATGAGAGAGCACTGTCTGCTGTGAAGAGAGAGCA 2798
Db 2589 AGAGTTTCTCGAAGAGAAACCAATTTAGTAGAGAGCGCTAGTCTGTGTGAAAAGAGCGGA 2648

```

Qy 2799 GAAAAATGGAGACAAACCTGAACTAAATTTGGAACAAACAAACGAGTATATACAGA 2858
Db 2649 GAAAAATGGAGACAAACCTGAACTAAATTTGGAACAAACAAACGAGTATATACAGA 2708
Qy 2859 GGCAGAGAGAGCTGGAGTGTATTTTGTAGATCTCAATATATATATATACAGCGGA 2918
Db 2709 GGCAGAGAGAGCTGGAGTGTATTTTGTAGATCTCAATATATATATATACAGCGGA 2768
Qy 2919 TACAAACATTTGGCATGATTCATCGGCGAGATATACTTGTTCATCGAATTCGAGAGGCTTA 2978
Db 2769 TACAGATATTCGCAATGATTCATCGGCGAGATATACTTGTTCATCGAATTCGAGAGGCTTA 2828
Qy 2979 TCTGTGAGAAATATCTGTATTCGCGGTGTAATTCGCGAAATTTTGAAGAAATTTAGAAGG 3038
Db 2829 TCTGCTGAGCTGTCTGTATTCGCGGTGTAATTCGCGAAATTTTGAAGAAATTTAGAAGG 2888
Qy 3039 TCGCATATATCACTGCAATCTCCCTATACAGATCGAGAGAAATGCTGTTAAATTTGTGATTT 3098
Db 2889 GCGTATTTTCACTGCAATCTCCCTATATGATCGAGAGAAATGCTATTAATAATTTGTGATTT 2948
Qy 3099 TAATAATGGATTTAGCATGCTGGAATGTAAAGAGGCGATGATGCT---ACAACAGAGCCA 3155
Db 2949 TAATAATGGCTTATCTGCTGGAACGTTGAAGGCGATGATGATGAGAGAAACAAACAA 3008
Qy 3156 TCACCGTTCTGCTGCTTGTATTCGAGAAATGGGAAGCAGAGTGTCAACAGCAGTTCCGCT 3215
Db 3009 CCAACGTTTCGCTGCTTGTATTCGGAATGGGAAGCAGAGTGTCAACAGAAAGTTCTGCT 3068
Qy 3216 CTGTCCGGGGGCTGCTATATCTTCGCTGTCAGAGCGTACAAAGAGGATTTGAGAGGG 3275
Db 3069 CTGTCCGGGGGCTGCTATATCTTCGCTGTCAGAGCGTACAAAGAGGATTTGAGAGGG 3128
Qy 3276 TTGTGTAACGATCCATGAAATCGAGAAACAAATACAGAGCAACTAAATTTTAAACACTGTA 3335
Db 3129 TTGTGTAACGATCCATGAAATCGAGAAACAAATACAGAGCAACTAAATTTTAAACACTGTA 3188
Qy 3336 AGAAGAGAGAGTGTATCCACGGGATACAGAACTGTAAATGATTTATCTGACACCAAGG 3395
Db 3189 AGAAGAGAGAGTGTATCCAAATAACAGGTAACGCTGAATGATTTATCTGTAATCAAGA 3248
Qy 3396 TACAGCAGTATGTAATTCGCTGATGCTGATGAGATGATGAGATGATGATGATGATGATG 3455
Db 3249 AGAATACGAGGTCGCT-----ACACTTCTCGTAATCGAGGATTAACGAAGCTCTTTC 3302
Qy 3456 AGCATCTGTTAAATCAACACGCTATGAAGAGAAACGATATACAGATGTACGAAGAGA 3515
Db 3303 CGTACAGCTGATTTATGCTGCTATGAGAGAAATCGTATACAGATGAGAGAGAGA 3362
Qy 3516 TAATCATTTGTAATGACAGAGGATGTAATTTATCCACCTACCACTGAGCTGTTATAT 3575
Db 3363 GAATCCTTGTGAAATTAACAGAGGATATAGGATTTACACGCTACCACTGAGTTGTTATGT 3422
Qy 3576 GACAAAGAAATTAAGATTTCCAGAAACCGGATAGGATGATGATGATGATGATGATGATG 3635
Db 3423 GACAAAGAAATTAAGATTTCCAGAAACCGGATAGGATGATGATGATGATGATGATGATG 3482
Qy 3636 GGAAGGGAAGTTTATGTTAGACAGCTGGAATTAATCTCTTATGAGGAATAG 3687
Db 3483 GGAAGGGAACATTTATGTTAGACAGCTGGAATTAATCTCTTATGAGGAATAG 3534

```

```

RESULT 13
US-10-365-645-25
; Sequence 25, Application US/10365645
; Publication No. US20030182682A1
; GENERAL INFORMATION: Thomas
; APPLICANT: Malvar, Amy Jelen
; TITLE OF INVENTION: Antibodies Immunologically Reactive with Broad-Spectrum
; TITLE OF INVENTION: Delta-Endotoxins (Amended)
; FILE REFERENCE: 11792, 0210.DVUS02 (MECO:210--3)
; CURRENT APPLICATION NUMBER: US/10365,645
; CURRENT FILING DATE: 2003-02-12

```

```

; PRIOR APPLICATION NUMBER: US 09/873,873
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 09/253,341
; PRIOR FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: US 08/922,505
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: US 08/754,490
; PRIOR FILING DATE: 1996-11-20
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Hybrid Delta-Endotoxin
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3531)
; US-10-365-645-25

Query Match      38.0%; Score 1402; DB 16; Length 3534;
Best Local Similarity 65.7%; Pred. No. 0;
Matches 2243; Conservative 0; Mismatches 1070; Indels 99; Gaps 10;

Qy 285 TCAGTGGGAAATTTCTCTAGAACATGTCGAAACAACTTATAAATCAACAAATAACAGAAAA 344
Db 213 TCAATGGGAGCATTTCTTGTACAAATTTGAACAGTAAATTAACCAAGATATAGAGATT 272
Qy 345 TGTAGGAATPACGGCACTTCTCTGATTAACAAGTTTAGGAGATTCCTTTAGAGCCTATCA 404
Db 273 CGCTAGGAACCAAGCCATTTCTAGATTTAGAAAGCAATCTTTATCAAAATTTACGC 332
Qy 405 ACAGTCACTTTGAGATTTGGCTAGAAACCGTATGATGCAAGAACGAGAGATGTTCTTTA 464
Db 333 AGAATCTTTTAGAGATTTGGGAGCAGATCTCTAATCCAGCATTAAGAAAGATGCG 392
Qy 465 TACCCAATATATAGCCTTTAGAACTTGATTTTCTTAATGCGATCGCGCTTTTCGCAATTAG 524
Db 393 TATTCAATTCATGACATGAAACAGTCGCCCTTACACCGCTATTTCTCTTTTGCAGTTCA 452
Qy 525 AAACCAAGAGATTTCAATTTAATGTTATATGCTCAAGCTGCAAAATTTACACCTATATT 584
Db 453 AAATTTATCAAGTTCTCTTTTATCAGTATATGTTTCAAGCTGCAAAATTTACATTTATCAGT 512
Qy 585 ATTGAGAGATGCTCTCTTTTTCGTAGTGAATTTGGCTTTACATCGCAGGAAATTTCAACG 644
Db 513 TTTGAGAGATGTTTTCAGTGTTTTGGACAAAGGTGGGATTTTGTATGCGCGACTATCAATAG 572
Qy 645 TTATTATGAGCGCAAGTGGAAACAAACGAGAGATTTATTCGACTATTTGCGTAGAATGGTA 704
Db 573 TCGTTATATGATTTAATAGGCTTATTTGCAACTATACAGATCACTGCTACGCTGTA 632
Qy 705 TAATACAGGCTCTAAATGATGTTGAGAGGACAAATGCGCGAAGTTGGTGGTGGTTATATCA 764
Db 633 CAATACGGGATTTAGAGCGTGTATGGGACCGGATTTCTAGAGATTTGGATAGATATATCA 692
Qy 765 ATTCCGTAGAGATCTAAAGTTAGGGGTATTTAGATCTAGTGGCACTATTTCCCAAGCTATGA 824
Db 693 ATTTAGAGAGAAATTAACACTAACTGTATTTAGATATGCTTTCTCTATTTCCGAACTATGA 752
Qy 825 CACTCGCACTTATCCCAATAAATACAGATGCTCAGTTTAAACAAAGGGAAGTTTATACAGACGC 884
Db 753 TAGTAGAACGTAATCCCAATTCGAACAGTTTCCCAATTTAAACAAAGAGAAATTTATACAAACCC 812
Qy 885 A--ATTGGAGCAACAGGGTAAATATATGCGAAGTATGAATTTGGTATTAATTAATGACCT 942
Db 813 AGTATTTAGAAAATTTTGTGTTAGTTTTCGAGCTCGGCTCAGGGCATAGAAAAGATAT 872
Qy 943 TCGTTTTCCGCTATAGAGACTGCGGTTATCCGAAGCCCGCATCTACTTGTATTTCTTAGAA 1002
Db 873 TAGAGTCCACATTTTATGATGATATATCTTAACAGTAAACCTTATACGAGTCTCATAG 932

```

QY 1003 CAACCTTACAAATTTTACGACTTTCATCAGATGGAGTGTACTAGGCATATGACTTACTGG 1062
Db 933 GGGTTATTATTATTTGGTCAGGGCATCA-----AATAATGGCTTCTCTGTAGGGTT 983
QY 1063 CGGGGGCACACAATTTCAATCTCGGCCCAATAGAGCGGGAATTAATACTTCAACGCATGGG 1122
Db 984 TTGGGGCCAGAAATTCATCTTTCCGCTATATGGAACTATGGGA-----ATGCA 1032
QY 1123 TCTACCAATATCTTCTATTAATTCCTGTGAAGATATCACTTCTCTCGAGAGCTATATGG 1182
Db 1033 GCTCCACAACAACCAATTTGTCTCAACTAGGTGAGGCGGTATAGAAACAATATATCGTCC 1092
QY 1183 ACTGAATCATATGAGGAGTCTTCTATGGGGAATTTACCTTGAACCTTATCATGGTGC 1242
Db 1093 ACTTTATAGAAGACCTTTTAATATAGGGATAAATAATCAACAACATCTCTGTTCTTGAC 1152
QY 1243 CCTACTGTAGATTTAATTTTAGAAACCTCAGAAATACTTTTGAAGAGGTACTGCTAAC 1302
Db 1153 GGGAC----AGAAATTCCTTATGGAACTCTCTCAAAATTTGCCATCCGCTGTATACAGAA 1208
QY 1303 TATAGTCAACCCCTATAGTCACTGGGCTTCAATTAAGAATTCAGAAACCTG-AAATPACC 1361
Db 1209 AAGCGGAACGCTAGATTCGCTGGATGAAATACCGCCACAGAAATAACAAGTGCACCTAG 1268
QY 1362 ACCGAAACACAGAACGACCAATTTATGAATCATATAGTATAGTATCTCACAATAGG 1421
Db 1269 GCAAGGAATTTAGTCAATCGATTAAGCCATGTTTCAATGTTTCGT-----TCAGGCTTTAG 1322
QY 1422 GCTCATTTCAAACTAGGTGGCATGTACAGATATTTCTTGGACGACCGTAGTGCAGA 1481
Db 1323 TAATAGTAGTATATATAGAGCTCCAAATGTTTCTTGGACGACCGTAGTGCAC 1382
QY 1482 TCGTACAAATACCAATTAGTTCAGATAGCATACAAATACCAATTCGTTGTAATAATCAATCAA 1541
Db 1383 CCCTACAAATACAAATGATCGGAGAGATTTACTCAAAATACCAATTCGTTGTAATAATCAATCAA 1442
QY 1542 CCTTAATTCAGGTACCTCTGTAGTCACTGAGGCGGAGGATTTACAGGAGGAGATATAATCCG 1601
Db 1443 ACTTCAGTCAGGTACTACTGTGTAAGAGGCGCGGGTTTACGGGAGGAGATATTTCTTCG 1502
QY 1602 AACTAACGTTAATGCTAGTGTACTAAGTATGGTCTTAATTTTAATAATACATCATTTACA 1661
Db 1503 ACGNACAAGTGGAGACCAATTCGCTTATACTATTTGTTTAATAATAATGGGCAATTTACCCCA 1562
QY 1662 GCGGTATCGGTGAGAGTTCGTTATGCTGCTTCTCAACAATGGTCTCGAGGGTAACTGT 1721
Db 1563 AAGGTATCGTGCAAGATACGCTATGCTCTACTACAAATCTAAGAAATTTACGTTAACGGT 1622
QY 1722 CGHAGGAGTACTACTTTTGATCAAGGATTCCTCTAGTACTATAGTGCAAAATGAGTCTTT 1781
Db 1623 TGCAGGTGAACGGATTTTGTGCTGCTCAATTTTAAACAAACAATGATACCGGTGACCCCAAT 1682
QY 1782 GACATCTCAATCATTTTAGATTTGAGAAATTTCTGTAGGTATTTAGTGCATCTGCGAGTCA 1841
Db 1683 AACATTCCAATCTTTTAGTACGCACTAATTAATACAGCTTTTATACATTCCCAATGAGCCA 1742
QY 1842 AACTGCT---GGAAATAGTAAAGTAAATGACAGGTAGACAAACGTTTTCACCTTTGATAA 1898
Db 1743 GAGTAGTTTCACAGTAGTCTGATACTTTTAGTTTCAGGGNAATGAAAGTTTATATAGACAG 1802
QY 1899 AATTGAATTCATTCCAATTAAGTCAACCTTCGGAAGAGAGATACGAATTTAGAAAGGGCGCA 1958
Db 1803 ATTTGAATTTGATTCAGTTACTGCAACATTTTGAAGCAGCAATATGATTTAGAAAGAGCACA 1862
QY 1959 AGAGCGGTGAATCTGTTTACTTAATACGAATTCGAAGAGATTTGAACACAGATGTGAC 2018
Db 1863 AAAGCGGTGAATGCGCTGTTTACTTCTATAACCAATAGGGATTAACCAAGATGTGAC 1922
QY 2019 AGATTATCATATTGATCAAGTATCCAAATTTAGTGGCGTGTGTTATCGGATGAATTTCTGCTT 2078
Db 1923 GGATTATCATATTGATCAAGTATCCAAATTTAGTGGATGTTTATCAGATGNAATTTTCTCT 1982
QY 2079 AGATGAAAGAGAGAAATTTACTTTGAGAAAGTGAATAATATGCGAAACGACTCAGTGTGTAAG 2138

Db 1983 GGATGAAAACGGAGAAATTTGTCGAGAAAGTCAAAACATTCGGAAGCGACTCAGTGTAGCGG 2042
QY 2139 AAACCTTACTCAAGATCCAAACTTTCATCATCCATCAATTAAGCAACCCAGACTTTCATATCTAC 2198
Db 2043 GAAATTTACTTCAAGATCCAAACTTTCAAAGGCATCAATAGGCAACTAGAC----- 2091
QY 2199 TAATGAGCAATCGAATTTTCATCTATCTCATGAACAATCTGGAACATGATGGTGGGGAAG 2258
Db 2092 -----CGTGGTTGGAGAGGAAG 2108
QY 2259 TGAGAACATTTACAAATCCAGGAAGGAAATGACGTATTTAAAGAGAAATTTACGTACACACTACC 2318
Db 2109 TAGCGATTTTACCAATCCAAAGAGAGAGATGACGTATTTCAAAGAAATTTATGTACACACTACC 2168
QY 2319 GGGGACCTTTTAATGAGTGTATCCGACGTATTTATATCAAAAAATAGGAGAGTCCGAATTT 2378
Db 2169 AGGTACCTTTTGTAGTGTCTATCAACATATTTGTATCAAAAAATCGATGAATCAAAATTT 2228
QY 2379 AAAAGCTTATATCTGCTACCAATTAAGAGGGTATATTTGAAGATAGTCAAGATTTAGAGAT 2438
Db 2229 AAAAGCCTTTTACCGCTTATCAATTTAAGAGGGTATATCGAAGATAGTCAAGACTTTAGAAAT 2288
QY 2439 ATATTTGATTCGTTTATAATCGAAACATGAAACATTCGATGTTTCCAGGTACCGAGTCCGT 2498
Db 2289 CTATTTAATTCGCTACAAATGCAAAACATGAACAGTAATTTGTCAGGTACGGGTTCTT 2348
QY 2499 ATGGCCGCTTTTCAAGTGTGAAGCCCAATCGGAAGGTGCGGAGAACCGAATCGATCGCACCC 2558
Db 2349 ATGGCCGCTTTTACGCCCAAAAGTCCAAATCGGAAGGTGTCGAGAGCCGAATCGATCGCGGCC 2408
QY 2559 ACATTTGAATGGAATCTGATCTAGATTTGTTCTCGAGAGATGGAGAAATATGTGCGCA 2618
Db 2409 ACACCTTTGAATGGAATCTGATTTAGATTTGTTCTGTTAGGGAATGGAGAAAGTGTGCCCCA 2468
QY 2619 TCATCCCATCATTTCTCTTTGGATATTTGATTTGGATGACACACTTTGCAATGAGAATCT 2678
Db 2469 TCATTCGATCATTTCTCTTTAGACATTTGATTTAGGATGTACAGACTTTAATAGGACCT 2528
QY 2679 AGGCGTGTGGTGTGATTTCAAGATTTAGACGCGAGGAAGGTCTATGCAAGACTTAGGGAATCT 2738
Db 2529 AGGTGTATGGTGTATCTTTAAGATTTAGACGCAAGATGGGCAACGACTTAGGGAATCT 2588
QY 2739 GGAATTTATTTGAAGAGAAACCAATTTAGAGAGAGCACTGTCTGCTGTGAAGAGAGCAGA 2798
Db 2589 AGAGTTTCTCGAAGAGAAACCAATTTAGTAGAGAGCGCTAGCTGCTGTGAAAAGAGCGGA 2648
QY 2799 GAAAATGAGAGAGACAAACGCTGAAAACCTACAAATTTGGAACAAACGAGTATATACAGA 2858
Db 2649 GAAAAAATGGAGAGACAAACGCTGAAAATTTGGAATGGGAACCAAAATATCGTTTATAAGA 2708
QY 2859 GGCAAAAGAGCTGTGGATGCTTTATTTGTAGATTTCTCAATATAATAGATTTACAAGCGGA 2918
Db 2709 GGCAAAAGATCTGTAGATGCTTTATTTGTAACTCTCAATATGATCAATTTACAAGCGGA 2768
QY 2919 TACAAACATTTGCAATGATTCATCGGAGATAAACTTTGTTTCATCGAATTCGAGAGGCTTA 2978
Db 2769 TACCAATTTTGCATGATTCATCGGAGATAAACTGTTTCATAGCATTCGAGAAGCTTA 2828
QY 2979 TCTGTCAAGATTTATCTGTTATCCGGGTGTAAATGCGGAATTTTTCGAAGAAATTTAGAGG 3038
Db 2829 TCTGCTGAGCTGTCTGTGATTCGGGTGTCAATTCGGGCTATTTTGAAGAAATTTAGAGG 2888
QY 3039 TCGCATTTTACATGCAATCTCCCTATACGATGCGAGAAATGTCTTAAAAATTTGTTGATTT 3098
Db 2889 GCGTATTTTACATGCAATCTCCCTATATGATGCGAGAAATGTCTATTAANAATTTGATTT 2948
QY 3099 TAATAATGGAATAGCATGCTGGAATGTAAAGGGCATGTAGATGT---ACAAAGAGCGCA 3155
Db 2949 TAATAATGGCTTATCTGCTGGAACGTTGAAGGGCATGTAGATGTAGAGAAACAAACAA 3008
QY 3156 TCACCGTTCCTCTGTTTATCCAGNAATGGGAAGCAGAAATGTCTCAAGCAGGTTCCGCT 3215

3009	CCAAAGTTCGGTCTCTTGTGTTCCGGAATGGGAAGCAAGAGTGTCACAAGAAGTTCGTGT	3068
3216	CTGTCCGGGGCGTGGCTATATCTCTCCGTGTACACAGCGGTACAAAGAGGGATATGAGAGGG	3275
3069	CTGTCCGGGTCGTGGCTATATCTCTTCTGTACACAGCGGTACAAGGAGGGATATCGAGAAGG	3128
3276	TTGTGTAACGATCCATGGAATCGAGAACAAATACAGAGCAACTAAAAATTTAAAAACTGTGA	3335
3129	TTGCGTAAACCATTCATGAGATCGAGAACAAATACAGACGAACTGAAGTTTATGCAACTCGCT	3188
3336	AGAAGGGAAGTGTATCCAAACGGATACAGAAACGTGTAATGATTATATCTGCACACCAAGG	3395
3189	AGAAGAGGAATCTATCCAAATAACACGGTAAACGTGTAATGATTATATCTGTAATCAAGA	3248
3396	TACAGCAGTATGTAAATTCGCCGTAAATGCTGGATATGAGGATGCATATGAAGTTGATCTAC	3455
3249	AGAAATACGAGGTCGCT-----ACACTTCTCGTAATCGAGGATATAACGAAGCTCCTTC	3302
3456	AGCATCTGTTAAATTACAAACCGACTTATGAAGAAGAAACGTATACAGATGTACGAAGAGA	3515
3303	CGTACCAGCTGATATATGCGTCAAGTCATGAAGAAATCGTATACAGATGGACGAAGAGA	3362
3516	TAATCATTTGGAATATGACAGAGGGTATGTGAATTATCCACCACCTACCACTCGGTTATAT	3575
3363	GAATCTCTTGTGAATTTAACAGAGGGTATAGGGATTTACACGCCACTACCACTTGGTTATGT	3422
3576	GACAAAGAATTAGAATACATTCCACAGAAAACCGATAAGGTATGGATTTCAGATTTCGAGAAAC	3635
3423	GACAAAGAATTAGAATACTTCCAGAAAACCGATATAGTATGATTTGAGATTTCGAGNAAC	3482
3636	GGAAAGGAAGTTTATTTGTAACACGCGTGGAAATTAATCTCTTTATCGAGGAATAG	3687
3483	GGAAAGAACATTTATCTGTGACACGCTGGAAATTAATCTCTTTATCGAGGAATAG	3534

RESULT 14

```

US-10-672-163-25
; Sequence 25, Application US/10672163
; Publication No. US20040093637A1
; GENERAL INFORMATION:
; APPLICANT: Malvar, Thomas
; APPLICANT: Gilmer, Amy Jelen
; TITLE OF INVENTION: Polynucleotide Compositions Encoding Broad Spectrum
; TITLE OF INVENTION: Delta-Endotoxins
; FILE REFERENCE: 11792.0215.DVUS02 MECO:2115--2
; CURRENT APPLICATION NUMBER: US/10/672,163
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: US 09/997,914
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 09/261,040
; PRIOR FILING DATE: 1999-03-02
; PRIOR APPLICATION NUMBER: US 08/754,490
; PRIOR FILING DATE: 1996-11-20
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Hybrid Delta-Endotoxin
; NAME/KEY: CDS
; LOCATION: (1)..(3531)
US-10-672-163-25

```

Qy	345	TGCTAGGAATAACGGCACTTGCTCGATTACAAAGTTTATAGGAGATTCCCTTTAGAGCGCTATCA	404
Db	273	CGCTAGGAACCAAGCCATTTCTGATTTAGAAGGCACTAAGCAATCTTTATCAAATTTACGC	332
Qy	405	ACAGTCACTTGAAGATTGGCTAGAAAAACCGTGATGATGCAAGACGAGMAGTGTTCTTTA	464
Db	333	AGAACTCTTTTAGAGAGTGGGAGCAGATCCCTACTAATCCAGCATTAAGAGAAGATGCG	392
Qy	465	TACCCAATATATAGCCTTAGAACTTGATTTTCTTAATTCGATGCGCTCTTTTCGAAATTAG	524
Db	393	TATTCAAATCCAATGACATGAACAGTCGCCCTTACAACCGCTATTCTCTCTTTTGCAGTTCA	452
Qy	525	AAACCAAGAGTTCATATTAATGGTATATGCTCAAGCTGCAAATTTACACCTTATATT	584
Db	453	AAATTAACAAGTTCCTCTTTTATCAGTATATGTTTCAAGCTGCAAATTTACATTTATCAGT	512
Qy	585	ATTGAGAGATGCCCTCTCTTTTGTGTAGTAATTTGGGCTTACATCGCAGGAAATTCACAG	644
Db	513	TTTGGAGAGATGTTTCAGTGTTTGGACAAAGGTGGGGATTTGATGCCCGGACTATCAATAG	572
Qy	645	TTATATAGGGCCCAAGTGGAAACAAACGAGAGATTAATTCGCACTATTTCGCTGAGAAATGGTA	704
Db	573	TCGTTAATAATGATTAACTAGGCTTAATGGCAACTATACAGATCATGCTGTGACGCTGGTA	632
Qy	705	TAATACAGGTCATAATAGCTTTGAGAGGACAAATTCGCCGAAGTTGGGTGGTGTATATATCA	764
Db	633	CAATACGGGATTAGAGCGTGTATGGGACCGGATTTAGAGATTTGGATAGATATAATATCA	692
Qy	765	ATTCGCTAGAGATCTAAACGTTAGGGGTATTAGATCTAGTGGCACTATTCCCAAGCTATGA	824
Db	693	ATTTAGAAGAGAAATTAACACTAACTGTATTAGATATCGTTTCTCTATTTCCGAACATGA	752
Qy	825	CACTCGCACTTATCCAAATAATACGAGTGTCTCAGTTAAACAGGGAAGTTTATACAGACGC	884
Db	753	TAGTAGAACGTATCCAAATTCGNAACGTTTCCCAATTAACAGAGAAATTTATACAAACCC	812
Qy	885	A-ATTGGAGCAACAGGGGTAAATATGGCAAGTATGAATTTGGTATATAATAATATGCACCT	942
Db	813	AGTATTAGAAAAATTTGATGGTAGTTTTCGAGGCTCGGCTCAGGGCATAGAAAGAAGTAT	872
Qy	943	TCGTTTTCCGCTATAGAGATCGCGGTTATCCGAAGCCCGCATCTACTTTGATTTTCTAGAA	1002
Db	873	TAGAGTCCCACTTTGATGGATATACTTAAACAGTATAACCATCTATACGGATGCTCATAG	932
Qy	1003	CAACTTACAAATTTTAGCACTTCATACGATGGAGTGTCTACTAGGCATATGACTTACTGG	1062
Db	933	GGGTTATTATTATTTGGTCAGGGATCA-----AATAATGGCTTCTCTGTAGGGTT	983
Qy	1063	CGGGGGCACAAATTCGAATCTCGGCCAATAGGAGCGGATTAATAATCCTCAACGCATGGG	1122
Db	984	TTCGGGCCAGAAATTCACCTTTCCGCTATATGAACTATGGGAA-----ATGCA	1032
Qy	1123	TCTACCAATACTTCTATTAAATCTGTAAAGATTATCATTTCTCTCGAGAGGTATATGG	1182
Db	1033	GTCTCCACAAACGATATGTTGTCTCAACTAGGTGAGGGCGGTATAGAACATTATCGTCC	1092
Qy	1183	ACTGAATCATATGCAGAGTGTCTTCTATGGGGAATTTACCTTCGAACCTATTCATGTGTCT	1242
Db	1093	ACTTTATATAGAGACCTTTTAAATATAGGGATAAATAATCAACAACTATCTGTCTTTGAC	1152
Qy	1243	CCTACTGTTATGATTTAAATTTTAGGAACCTCTCAGATACTTTTGAAGAGAGTACTGCTAAC	1302
Db	1153	GGGAC----AGAAATTTGCTTATGGAACTCTCTCAAAATTTGCCATCGCTGTATACAGAAA	1208
Qy	1303	TATAGTCAACCCCTATGATGTCACTCGGCTCTCAATTTAAAGATTTCAGAAAACTG-AAATTACC	1361
Db	1209	AAGCGNAACGGTAGATTTCGTGATGAATAATACCGCCACAGAAATAACACGTGCCACCTAG	1268
Qy	1362	ACCAGAAACAAACAGAACGCAAAATTAATGAAATCATATAGTCATAGGTTATCTCACATAGG	1421
Db	1269	GCNAAGATTTAGTCATTCGATTAAGACCATGTTTCAAAGTTTCGT-----TCAGGCTTTAG	1322

Db 3423 GACAAAAGAAATTAGATACTTCCAGAAACCCGATAGGTATGAGATTGAGAAAC 3482

Qy 3636 GGAAGGGAAGTTTATTGTAGACAGCGTGGAACTACTCTTATGAGGAATAG 3687

Db 3483 GGAAGGAACATTTATCGTGGACGCGTGGAAATTAATCTCTTATGAGGAATAG 3534

RESULT 15

US-10-739-482-25

Sequence 25, Application US/10739482

Publication No. US20040132975A1

GENERAL INFORMATION:

APPLICANT: Malvar, Thomas

APPLICANT: Mohan, Komarlingham S.

APPLICANT: Sivapramaniam, Sakuntala

TITLE OF INVENTION: Polynucleotide Compositions Encoding Broad Spectrum

TITLE OF INVENTION: Delta-Endotoxins

FILE REFERENCE: MECO:220--1

CURRENT FILING DATE: 2003-12-18

PRIOR APPLICATION NUMBER: US/10739,482

PRIOR FILING DATE: 2000-08-11

PRIOR APPLICATION NUMBER: US 6,242,241

PRIOR FILING DATE: 1999-02-19

PRIOR APPLICATION NUMBER: US 6,110,464

PRIOR FILING DATE: 1997-09-03

PRIOR APPLICATION NUMBER: US 6,017,534

PRIOR FILING DATE: 1996-11-20

NUMBER OF SEQ ID NOS: 35

SOFTWARE: Patent in version 3.2

SEQ ID NO 25

LENGTH: 3534

TYPE: DNA

ORGANISM: Artificial sequence

FEATURE:

OTHER INFORMATION: Hybrid Delta-Endotoxin

NAME/KEY: CDS

LOCATION: (1)...(3531)

US-10-739-482-25

Query Match 38.0%; Score 1402; DB 18; Length 3534;

Best Local Similarity 65.7%; Pred. No. 0;

Matches 2243; Conservative 0; Mismatches 1070; Indels 99; Gaps 10;

Qy 285 TCAGTGGGAATTTCTCTAGACATGTGCGAACACTTATAAATCAACAATACAGAAA 344

Db 213 TCAATGGGACGCAATTTCTGTACAAATGGAACAGTTAATTAACCAAGAAATAGAAAT 272

Qy 345 TGCTAGGAATACGGCACTTCTCGATTACAAAGGTTTAGGAGATTCTCTTAGAGCCTATCA 404

Db 273 CGCTAGGAACCAAGCCATTTCTAGATTAGAGGACTAAGCAATCTTTATCAATTTAGCG 332

Qy 405 ACAGTCACTTGAAGATTGGCTAGAAAACCGTGATGATGCAAGAACGAGAAAGTGTCTTTA 464

Db 333 AGAATCTTTTAGAGAGTGGGAAGCAGATCCTACTAATCCAGCATTAAGAGAAAGAGATGCG 392

Qy 465 TACCAATATAGCCCTTAGAATCTGATTTCTTAATGCGATGCCGCTTTTCGCAATTAG 524

Db 393 TATTCAATCAATGACATGAACAGTGCCTTTACAAACCGCTATTCTCTTTTGGAGTTCA 452

Qy 525 AAACCAAGAAAGTTCATTAATAGTGTATATGCTCAAGCTGCAAAATTTACACCTATTATT 584

Db 453 AAATTATCAAGTCTCTTTTATCAGTATATGTTCAAGCTGCAAAATTTACATTTACAGT 512

Qy 585 ATTGAGAGATGCTCTCTTTTGGTAGTAATTTGGGCTTACATCGCAGAAATTTCAACG 644

Db 513 TTTGAGAGATGTTTCACTGTTTGGACAAAGTGGGATTTGATCGCCGCACTATCAATAG 572

Qy 645 TTATTATGAGCGCAAGTGGAAACAAACGAGAGATTATTCCGACTATTTCGCTAGAAATGGTA 704

Db 573 TCGTTATATGATTTAACTAGGCTTATTGGCAACTATACAGATCATCTGCTGACGCTGTTA 632

Qy 705 TAATACAGGCTTAAATAGCTTTGAGAGGGAACAAATGCCGAAGTTGGGTGGTGTATATCA 764

Db 633 CAATACGGGATTAGACGCTGTATGGGACCGGATTTAGAGATTGGATAGATATATCA 692

Qy 765 ATTCCGCTAGAGATCTTAACGTTAGGGGTATTAGTCTAGTGGCACTATTCCCAAGCTAAGA 824

Db 693 ATTTAGAAGAGAATTAACACTAATCTGTATTAGATATCGTTTCTCTATTTCCGAACATGA 752

Qy 825 CACTCGCACTTATCCAAATAATACGAGTGTCTAGTTAAACAGGGAAGTTTATACAGCGC 884

Db 753 TAGTAGAACGTTATCCAAATTCGAACAGTTTCCAAATTAACAAGAGAATTTATACAAACC 812

Qy 885 A--ATTGGAGCAACAGCGGTAAATATATGGCAAGTATGAATTTGGTATATAATAATATGACCT 942

Db 813 AGTATTAGAAAATTTTGATGGTAGTTTTCGAGGCTCGGCTCAGGCGCATAGAAAGATAT 872

Qy 943 TCGTTTTCGCTATAGAGACTGCGGTTATCCGAAGCCCGCATCTACTTTGATTTTCTAGAA 1002

Db 873 TAGGAGTCCACATTTGATGGATATCTTAACAGTATAACCATCTATACGGATGCTCATAG 932

Qy 1003 CAATTTACAATTTTAGCACTTTCATCAGATGGAGTGTCTACTAGGCATATGACTTACTGG 1062

Db 933 GGGTTATTATTATTGGTCAGGGCATCA-----ATAATGGCTTCTCTCTGTAGGGTT 983

Qy 1063 CGGGGCGACACAATCAATCTCGGCAATAGGAGGGGATTAATACTCAACGCGATGGG 1122

Db 984 TTCGGGGCCAGAATTCACITTTCCGCTATATGAACTATGGGA-----ATGCA 1032

Qy 1123 TCTACCAATACITTTCTATTAATCTCTGAAGATATCATCTTCTCTCGAGACGATATATGG 1182

Db 1033 GCTCCACAACAACGTTATTGTTGCTCAACTAGGTGAGGGCGGTATAGAACATTTATCGTCC 1092

Qy 1183 ACTGAATCATATGCGAGAGTGTCTTATGGGAATTTACTTGAACCTATTCATGTTGTC 1242

Db 1093 ACTTTATATAGAAAGCCTTTTAATAAGGATAATAATAACAACATCTGTTCTTTGAC 1152

Qy 1243 CCTACTGTTAGATTTAAATTTTAGAAACCTCTCAGAACTATTTTGAAGAGAGTACTGCTAAC 1302

Db 1153 GGGAC---AGATTTGCTTATGGAACCTCTCAATTTGCCATCGCTGTATACAGAA 1208

Qy 1303 TATAGTCAACCCCTATGAGTCACTGGGCTTCAATTTAAAGATTCAGAACTG--AATTACC 1361

Db 1209 AAGCGGAACGGTAGATTTCGTGTGATGAAATACCGCACAGAAATAACAGTGCACCTAG 1268

Qy 1362 ACCAGAAACAACAGAACGACCAAAATTAAGAACTATATAGTATAGTGTATCTCACATAGG 1421

Db 1269 GCAAGATTATTGATCTGATTAAGCCATGTTTCAATGTTTCGT-----TCAGCTTTAG 1322

Qy 1422 GCTCATTTCAATCTAGGTTGATGTACAGTATATTTTGGACGACCGTAGTGCAGA 1481

Db 1323 TAATAGTAGTGTAAATTAATAGAGCTCCAATGTTTCTTGGACGACCGTAGTGCAAC 1382

Qy 1482 TCGTACAAATACATTTAGTTTCAGATAGATACACAAATACCATTTGGTAAATTCATTCAA 1541

Db 1383 CCCTACAAATACAAATTTGATCCCGAGAGAGATTACTCAATACCATTTGGTAAACACATAC 1442

Qy 1542 CCTTAATTCAGGTACCTCTGTAGTCACTGAGTCCCGAGGATTTTACAGGAGGAGATATATCCG 1601

Db 1443 ACTTCAGTCAAGTACTGTTGTTAGAGAGCCCGGGTTTACGGGAGGAGATATCTTCG 1502

Qy 1602 AACTAAAGTTAATGGTGTACTTAAGTAGTGGTCTTAAATTTTAAATAATACATCATTA 1661

Db 1503 ACGAAACAGTGGAGGACCATTTGCTTATACTATTGTTAATAATAAATGGGCAATACCCCA 1562

Qy 1662 GCGGTATCGGTGAGAGTTCGTTATGCTGTTCTCAACAAATGGTCTCGAGGAGTACTGT 1721

Db 1563 AAGGTATCGTGCAGAAATAACGTTATGCTTACTACTACAAATCTAAGAAATTTACGTAACGGT 1622

Qy 1722 CGGAGGAGTACTTACTTGTGATCAAGATTTCCCTAGTACTATGAGTGCAGAAATGAGTCTTT 1781

Db 1623 TGCAGGTGAACGGATTTTGTGCTGCTCAATTTAACAACAAATGGATACCGGTGACCCAT 1682

Qy 1782 GACATCTCAATCATTTAGATTTGAGAAATTTCTGTAGGTATTAGTGCATCTGCGCAGTCA 1841

[illegible]

2709	GGCAAAAGAAATCTCTAGATGCTTTATTTGTAAACTCTCAATATGATCAAAATTACAAGCGGA	2768
2919	TACAAACATTTGGCATGATTCATCGGCAGATAAACTTGTTCATCGAATTCGAGAGGCTTA	2978
2769	TACGATATTTGCCATGATTCATCGGCAGATAAACGTGTTTCATAGCATTTCGAGAGCTTA	2828
2979	TCTGTCAAGAAATATCTGTATCCCGGTGTAAATGCGGAAATTTTGAAGAAATTAGAAGG	3038
2829	TCTGCTGAGCTGCTGTGATTCGGGTGTCAATCGGCTATTTTGAAGAAATTAGAAGG	2888
3039	TCGCATTTACATCGCAATCTCCCTATACGATGCGAGAAATGTGTTTAAATAATGGTGATTT	3098
2889	CGGTATTTTCACTGCATCTCCCTATATGATGCGAGAAATGTCAATTTAAATAATGGTGATTT	2948
3099	TAATAATGGATAGCATGCTGGATGTAAAGGGCATGTAGATGT---ACAAACAGAGCCA	3155
2949	TAATAATGGCTATCTCTGCTGGAAACGTGAAGGGCATGTAGATGTAGAGAAACAAACAA	3008
3156	TCACGTTCTGTCTCTGTTATCCAGAAATGGGAAGCAGAAAGTGTCAACAAGCAGTTCGCGT	3215
3009	CCAACTTCGCTCGTCTGTTGTCGGAAATGGGAAGCAGAAAGTGTCAAGAGATTTGTTGT	3068
3216	CTGTCCGGGGCGTGGCTATATCTCTCGTGTACAGCGTACAAAGAGGATATGGAGAGG	3275
3069	CTGTCCGGGTGCTGGCTATATCTTCTGTGTACAGCGTACAAAGAGGATATGGAGAGG	3128
3276	TTGTGTAAACGATCCATGAATTCGAGAAACAATACAGACGAACTAAAAATTTAAAACTGTGA	3335
3129	TTGCTTAACCATTCATGAGATCGAGAACAAATACAGACGAACTGAAGTTTAGCAACTGCGT	3188
3336	AGAAGAGAGAGTGTATCCACGGATACAGAAAGTGTATGATTTATCTCCACACCAAGG	3395
3189	AGAAGAGGAAATCTATCCAAATAACACGGTAAACGTGTAAATGATTACTGTAATCAAGA	3248
3396	TACAGCAGTATGTAATTTCCCGTAATGCTGGATATGAGGATGCATATGAAGTTGATCTAC	3455
3249	AGAATACGGAGGTGCGT-----ACACTTCTCGTAATTCGAGGATATAACGAAGCTCTCTC	3302
3456	AGCATCTGTTAATTAACAAACCGACTATGAAGAAGAAACGTTATACAGATGTACGAAGAGA	3515
3303	CGTACACAGCTGATTTATCGGTAGTCTATGAAGAAATAATCGTTATACAGATGGACGAAGAGA	3362
3516	TAATCATTTGTAATATACACAGAGGATGTGTAATTTATCCACCCTACCACTCGTGTATAT	3575
3363	GAATCCTTTGTGAATTTAACAGAGGATATAGGATTTACACGCCACTACCACTGTTATGT	3422
3576	GACAAAAGAAATAGAAATCTTCCAGAAACCGATAAAGGATATGATTGAGATTCGAGAAAC	3635
3423	GACAAAAGAAATAGAAATCTTCCAGAAACCGATNAAGGTATGATTGAGATTCGAGAAAC	3482
3636	GGAAAGGAAGTTTATTTGTAGACAGCGTGAATTTACTCTCTTATCGAGGAATAG	3687
3483	GGAAAGGAACATTTATCTGTGACAGCGTGAATTTACTCTTATCGAGGAATAG	3534

Search completed: April 21, 2005, 04:10:22
Job time : 1965 secs

This Page Blank (uspto)

Db 557 TMEIGENLTSRTFRYTDFSNPFSPFRANPDIIGISEQPLFGAG-SISG-----ELYIDKIE 611
Qy 636 FIPITATFEAEYDLERAQAVNALFTNTNPRRLKTDVTDYHIDQVSNLVACLSDSEFCCLDE 695
Db 612 IILADATFEAEYDLERAQAVNALFTSSNOIGLKTVDYHIDQVSNLVCLSDSEFCCLDE 671
Qy 696 KRELLEKVKYAKRLSDERNLLQDPNFTSINKQDPFISTNEQSNFTSIHQSEHGHWGSEN 755
Db 672 KRELSEKVKHAKRLSDERNLLQDPNFRGINRQPD-----RGRGSTD 713
Qy 756 ITIQGNDVPEKENYVTLPGTFNECYPTLYLQKIGESLKAATRYOLRGYIEDSDLEIYL 815
Db 714 ITIQGDDVPEKENYVTLPGTFNECYPTLYLQKIDESKLKAYTRYELRGYIEDSDLEIYL 773
Qy 816 IRYNAKHETLDVPGTESVPLSVESPIGRGCEPNRCAPHPFEMNPDLDCSCRDGEKCAHHS 875
Db 774 IRYNAKHETLVNPGTSLNPLSAQSPIGKCGEPNRCAPHPLEWNPDLDCSCRDGEKCAHHS 833
Qy 876 HHFSLDIDIGCTDLHENLGVVVFVKITQEGHARLGNLEFIEBKPLLGEALSRYKGAEEK 935
Db 834 HHFTLDIDVGCTDLNEDLGVVVFVKITQEGHARLGNLEFIEBKPLLGEALSRYKGAEEK 893
Qy 936 WRDKREKLOLETXRVYTEAKEAVDALFVDSOYNRLQADTNIQMIHAADKLAVHRIREAYLS 995
Db 894 WRDKREKLOLETNIVYKEAKESVDALFVNSOYDLQVDTNIAIHAADKRVHRIREAYLP 953
Qy 996 ELSVIPGVNAEIPFEELEGRITITSLYDARNVVKNGDFNNGLACWNVKGHVY-QQSHHR 1054
Db 954 ELSVIPGVNAEIPFEELEGRITITAYSLYDARNVVKNGDFNNGLACWNVKGHVYBQQNHR 1013
Qy 1055 SVLVIPWEAEVSAQVRVCPGRGYILRVTAIYKEGYGEGCVTIHEIENNTDELKFKCEE 1114
Db 1014 SVLVIPWEAEVSAQVRVCPGRGYILRVTAIYKEGYGEGCVTIHEIENNTDELKFKCEE 1073
Qy 1115 EVYPTDTGTCNDYTA-----HQTAVCMNRNAGYEDAYEDVDTTASVNYKPTYBEETDVR 1170
Db 1074 EVYFNNTVTCCNYTGTQEEYEGT--YTSRNGQYDEAYGNNPSVPADYASVYEEKSVTDGR 1131
Qy 1171 RDNHCEYDRGVNYPPLPAGVMTKELFPETDKWIEIGETGKFTVDSVELLMEE 1228
Db 1132 RENPCSNRGYDYPPLPAGVMTKELFPETDKWIEIGETGKFTVDSVELLMEE 1189

Search completed: April 21, 2005, 04:43:58
Job time : 1519 secs

```
Db 100 EGLGNFNIVYAEFKWEEDNNPATRTRVIDRFRILDGLLERDIPSRISGFVFPVLLSV 159
Qy 185 YAAANLHLLLRDASLFGSEGLTSQEIORYYERQVETRDYSDYCVWYNTGLNSLRG 244
Db 160 YAAANLHLLLRDASLFGSEGLTSQEIORYYERQVETRDYSDYCVWYNTGLNSLRG 219
Qy 245 TNAASWRYNQFRDLTLGVLDLVALPSPYTRTYPIINTSAQLTRVYVTAIGATGNMA 304
Db 220 STYQDWITYNLRDLTLVLDIAAFPNYDNRYPYQPVQQLTRVYVTDPL----INFN 275
Qy 305 SMWYNNAPSFAIETAVIRSPHLLDFLEQLTIFSTSSRSWATRHMTYWRGHTIQSRPI 364
Db 276 PQLOSVAQLPTFNWMESSAIRNPHLFDILNLTFTD---WFSVGRNFYWGGRVSSLI 332
Qy 365 GGLNTSTHGSTNTSINPRLSFRSDVYWTESYAGVLL----WGIVLEPHIGVPTVRFN 420
Db 333 GGGNITSPYIGREANQEPSPFTFNGPVFRTLSPNPLRLQLQPPAPPNLRCVGEVFS 392
Qy 421 FRNPQNTFE---RGTYNSQPYESPGQLKDSSETLPPTETTERPNYVESHRLSHIGLIS 477
Db 393 --TPTNSFTYRGRTV-----DSLTELPPEDNSVPPREGYSHRLCHATFVQ 436
Qy 478 QSRVHV----PVYSWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVWSGPGFTGDI 533
Db 437 RSGTPFLTGTGVFSWTHRSATLTNTIDPERINQIPLVKGFVWGGSVTITGPGFTGDI 496
Qy 534 RTNVGSLVSMGLNFNTSILQRYRVRVYAAO-----TWLVRVTVGGSTTFDQGFPS 586
Db 497 RRTTFGDFVSLQVNSIPITQRYRLFRYASSRDARVILVTGAASTGVGGVSNVMPLOK 556
Qy 587 TMSANESLTSQSFRFAFP-----VGISASGSQTAGISISNNAGRTQTFHDKIE 635
Db 557 TMEIGENLTSKTRFYDFSPPFRANPDIIIGISEQLFCAG-SISSG-----ELYIDKIE 611
Qy 636 FIPITAFEAAYDLERAQEAVALFTWNPRLKTDVTDHIDOVSNLVACLDSDEFCLDE 695
Db 612 IILADATFEASDLERAQKAVNALFTSSNQIGLTKTDVTDHIDQVSNLVDCLDSDEFCLDE 671
Qy 696 KRELLEKVKAKRLSDERNLQDPNFTSINKQDPFISTNEQSNFTSITHEQSEHWGWSN 755
Db 672 KRELSEKVKAKRLSDERNLQDPNFTSINKQDPFISTNEQSNFTSITHEQSEHWGWSN 713
Qy 756 ITIQGNDVFKENYVTLPGTFNECYPTLYQKIGESLKYATRYQLRGYIETDSQDLIYL 815
Db 714 ITIQGNDVFKENYVTLPGTFNECYPTLYQKIGESLKYATRYQLRGYIETDSQDLIYL 773
Qy 816 IRYNAKHETLDVCTESVWPLSVESPTGRCGNPCAPHEFWNPDLDCSCRDGKCAHHS 875
Db 774 IRYNAKHETLDVCTESVWPLSVESPTGRCGNPCAPHEFWNPDLDCSCRDGKCAHHS 833
Qy 876 HIFSLDIDIGCTDLHENLGVVWFKIQEGHARLGNLEFIEEKPLLGEALSRVKAERK 935
Db 834 HHTFLDIDVCTDLNEDLGVVWFKIQEGHARLGNLEFIEEKPLLGEALSRVKAERK 893
Qy 936 WRDKREKQLQLETKRVYVTEAKVADALFVDSQYRNLQADTNIGMTHAADKLVRHIREAYLS 995
Db 894 WRDKREKQLQLETKRVYVTEAKVADALFVDSQYRNLQADTNIGMTHAADKLVRHIREAYLS 953
Qy 996 ELSVIPGVNAIPEELEGRITIALSYDARNVKNQGFNNGLACWNVKGHVDV-QQSHR 1054
Db 954 ELSVIPGVNAIPEELEGRITIALSYDARNVKNQGFNNGLACWNVKGHVDV-QQSHR 1013
Qy 1055 SVLVIPWEAEVSAQVRCVGRGYILRVATYKGYGEGCVTIHEIENTDELKFNCEBE 1114
Db 1014 SVLVIPWEAEVSAQVRCVGRGYILRVATYKGYGEGCVTIHEIENTDELKFNCEBE 1073
Qy 1115 EYVPTDGTGNDYTA----HQTAVCNRSNAGYEDAYEDVDTASVNYKPYEBEYTDVR 1170
Db 1074 EYVPTDGTGNDYTA----HQTAVCNRSNAGYEDAYEDVDTASVNYKPYEBEYTDVR 1131
Qy 1171 RDHCEYDRGVNYPPLPAGYMTKELEYFPETDKWIEIGETEGKFIVDSVELLMEE 1228
Db 1132 RENPCSNRGYGDYTPLPAGYMTKELEYFPETDKWIEIGETEGKFIVDSVELLMEE 1189
```

RESULT 15

```
US-10-926-819-9
; Sequence 9, Application US/10926819
; Publication No. US20050049410A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-003, A Delta-Endotoxin Gene and
; TITLE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/281577
; CURRENT APPLICATION NUMBER: US/10/926,819
; CURRENT FILING DATE: 2004-08-26
; PRIOR APPLICATION NUMBER: 60/498,518
; PRIOR FILING DATE: 2003-08-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-926-819-9

Query Match . 54.0%; Score 3495.5; DB 17; Length 1189;
Best Local Similarity 56.9%; Pred. No. 6.9e-259;
Matches 716; Conservative 143; Mismatches 290; Indels 109; Gaps 20;

Qy 7 NENEIT--NALSIIPAVSNHSTOMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGNIAG 64
Db 5 NQNCIPYCNLS-----NPEVLLDGERISTGN-----SSIDISLSLVQ 43
Qy 65 RIILGVIGVPPAGQIAFYSLVGLMWRGRDQWEI FLEHVEQLINQOITNANRTALRL 124
Db 44 FLVSNF-VPGGGFLVGLIDFVWGIVGP---SOWDAFLVQIEQLINERIAEFARNAAILN 99
Qy 125 QGIGDSFRAQQOLESLEWLENRDARTSRVLYTOYIALELDFLNAMPFAIRNOEVPLLMV 184
Db 100 EGGNNFNIVYAEFKWEEDNNPATRTRVIDRFRILDGLLERDIPSRISGFVFPVLLSV 159
Qy 185 YAAANLHLLLRDASLFGSEGLTSQEIORYYERQVETRDYSDYCVWYNTGLNSLRG 244
Db 160 YAAANLHLLLRDASLFGSEGLTSQEIORYYERQVETRDYSDYCVWYNTGLNSLRG 219
Qy 245 TNAASWRYNQFRDLTLGVLDLVALPSPYTRTYPIINTSAQLTRVYVTAIGATGNMA 304
Db 220 STYQDWITYNLRDLTLVLDIAAFPNYDNRYPYQPVQQLTRVYVTDPL----INFN 275
Qy 305 SMWYNNAPSFAIETAVIRSPHLLDFLEQLTIFSTSSRSWATRHMTYWRGHTIQSRPI 364
Db 276 PQLOSVAQLPTFNWMESSAIRNPHLFDILNLTFTD---WFSVGRNFYWGGRVSSLI 332
Qy 365 GGLNTSTHGSTNTSINPRLSFRSDVYWTESYAGVLL----WGIVLEPHIGVPTVRFN 420
Db 333 GGGNITSPYIGREANQEPSPFTFNGPVFRTLSPNPLRLQLQPPAPPNLRCVGEVFS 392
Qy 421 FRNPQNTFE---RGTYNSQPYESPGQLKDSSETLPPTETTERPNYVESHRLSHIGLIS 477
Db 393 --TPTNSFTYRGRTV-----DSLTELPPEDNSVPPREGYSHRLCHATFVQ 436
Qy 478 QSRVHV----PVYSWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVWSGPGFTGDI 533
Db 437 RSGTPFLTGTGVFSWTHRSATLTNTIDPERINQIPLVKGFVWGGSVTITGPGFTGDI 496
Qy 534 RTNVGSLVSMGLNFNTSILQRYRVRVYAAO-----TWLVRVTVGGSTTFDQGFPS 586
Db 497 RRTTFGDFVSLQVNSIPITQRYRLFRYASSRDARVILVTGAASTGVGGVSNVMPLOK 556
Qy 587 TMSANESLTSQSFRFAFP-----VGISASGSQTAGISISNNAGRTQTFHDKIE 635
Db 557 TMEIGENLTSKTRFYDFSPPFRANPDIIIGISEQLFCAG-SISSG-----ELYIDKIE 611
Qy 636 FIPITAFEAAYDLERAQEAVALFTWNPRLKTDVTDHIDOVSNLVACLDSDEFCLDE 695
Db 612 IILADATFEASDLERAQKAVNALFTSSNQIGLTKTDVTDHIDQVSNLVDCLDSDEFCLDE 671
Qy 696 KRELLEKVKAKRLSDERNLQDPNFTSINKQDPFISTNEQSNFTSITHEQSEHWGWSN 755
Db 672 KRELSEKVKAKRLSDERNLQDPNFTSINKQDPFISTNEQSNFTSITHEQSEHWGWSN 713
Qy 756 ITIQGNDVFKENYVTLPGTFNECYPTLYQKIGESLKYATRYQLRGYIETDSQDLIYL 815
Db 714 ITIQGNDVFKENYVTLPGTFNECYPTLYQKIGESLKYATRYQLRGYIETDSQDLIYL 773
Qy 816 IRYNAKHETLDVCTESVWPLSVESPTGRCGNPCAPHEFWNPDLDCSCRDGKCAHHS 875
Db 774 IRYNAKHETLDVCTESVWPLSVESPTGRCGNPCAPHEFWNPDLDCSCRDGKCAHHS 833
Qy 876 HIFSLDIDIGCTDLHENLGVVWFKIQEGHARLGNLEFIEEKPLLGEALSRVKAERK 935
Db 834 HHTFLDIDVCTDLNEDLGVVWFKIQEGHARLGNLEFIEEKPLLGEALSRVKAERK 893
Qy 936 WRDKREKQLQLETKRVYVTEAKVADALFVDSQYRNLQADTNIGMTHAADKLVRHIREAYLS 995
Db 894 WRDKREKQLQLETKRVYVTEAKVADALFVDSQYRNLQADTNIGMTHAADKLVRHIREAYLS 953
Qy 996 ELSVIPGVNAIPEELEGRITIALSYDARNVKNQGFNNGLACWNVKGHVDV-QQSHR 1054
Db 954 ELSVIPGVNAIPEELEGRITIALSYDARNVKNQGFNNGLACWNVKGHVDV-QQSHR 1013
Qy 1055 SVLVIPWEAEVSAQVRCVGRGYILRVATYKGYGEGCVTIHEIENTDELKFNCEBE 1114
Db 1014 SVLVIPWEAEVSAQVRCVGRGYILRVATYKGYGEGCVTIHEIENTDELKFNCEBE 1073
Qy 1115 EYVPTDGTGNDYTA----HQTAVCNRSNAGYEDAYEDVDTASVNYKPYEBEYTDVR 1170
Db 1074 EYVPTDGTGNDYTA----HQTAVCNRSNAGYEDAYEDVDTASVNYKPYEBEYTDVR 1131
Qy 1171 RDHCEYDRGVNYPPLPAGYMTKELEYFPETDKWIEIGETEGKFIVDSVELLMEE 1228
Db 1132 RENPCSNRGYGDYTPLPAGYMTKELEYFPETDKWIEIGETEGKFIVDSVELLMEE 1189
```

```

; CURRENT APPLICATION NUMBER: US/10/200,522
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/337,280
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 08/980,071
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 08/757,536
; PRIOR FILING DATE: 1996-11-27
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 61
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant Delta Endotoxin
US-10-200-522-61

Query Match          54.0%; Score 3496.5; DB 14; Length 1189;
Best Local Similarity 56.9%; Pred. No. 5.7e-259;
Matches 716; Conservative 143; Mismatches 290; Indels 109; Gaps 20;

QY 7 NENEII--NALSIAPVNSHSTQMDLSPDARIEDSLCTIAEGNNINPLVSASTVQTGINIAG 64
DB 5 NQNCIPYNCLS-----NPEVLLDGERISTGN-----SSIDISLSLVQ 43

QY 65 RILGLVGPVAGQIASFYSFLVGLMWRGRDQWEIFLEHVEQLINOQITENARNTALARL 124
DB 44 FLVSNF-VPGGFLVGLIDFWGIVGP---SQWDAFLVQIEQLINERIAEFARNAAIL 99

QY 125 QGLGDSFRAYQOQSLDMLNRRDDARTSVLYTQVIALELDFLNAFLPFAIRNOEVLMMV 184
DB 100 EGLGNPNIVYEAKEWEDPNPATRVIDRFRILDGLLDRDIPSDISGPEVPLLSV 159

QY 185 YAAANLHLLLRDASLFGSEFGLTSEIQRYERQVEQTRDYSDYCVEMYNGLSLRG 244
DB 160 YAAANLHLLLRDASLFGSEFGLTSEIQRYERQVEQTRDYSDYCVEMYNGLSLRG 219

QY 245 TNAASVRYNQFRDLPLGLDLVALPSPYDTRTPINTSAQLTRVYTDALGATGVNMA 304
DB 220 STYQDMITYNLRDLTLVLDIAAFPNDYRNPPIQPVQGLTRVYTDPL---INFN 275

QY 305 SMWYNNAPSAFAETAVIRSHPLDLFLBOLITFTSSRWASATRMVYWRGHTIQSRPI 364
DB 276 POLQSAVLAQLETFNWMSSAIRNPHLFDLNLATFTD---WFSVGRNFYWGGRVSSLI 332

QY 365 GGLNTSTHGSNTSINPVLSPFSRDVYTESVAGVLL---WGIYLEPIHGVPVTVREN 420
DB 333 GGGNITSPIYGREANQEPFRSFTFNGPVFRTLNSPTLRLIQQPWPAPPFNLRGVEGVEFS 392

QY 421 FRNPQTFE---RGTYNSQPYSPGLQLKDSSETLPPETTERPNYSESYSHRLSHIGLIS 477
DB 393 --TPTNSFTYRGRTV-----DSLTELPEPDSVPPREGYSHRLCHATFVQ 436

QY 478 QSRVHV---PVYSWTHRSADRTNITSDDSIQTIPLVKSNLNGSTSVSGPFTGDDII 533
DB 437 RSGTFFLTGTVFSTWTHRSATLNTIDPERINQIPLKGFVWGGTSTVITGPGFTGDDIL 496

QY 534 RTVNGSVLGNLFNNTSLQRYVRVRYAASQ-----TMVLRTVTVGGSTTFDQGFPS 586
DB 497 RRNTFGDFVSLQVNSINPITQRYRLFRYASSRDARVILVTGAASTGVGGVSNVMPLOK 556

QY 587 TMSANESLTSQSPFAFPF-----VGTISASGQTAGISINNACRQFPHFKIE 635
DB 557 TMEIGENLRTFRYTFDPSFPFRANDPIIGISEQLPFGAG-SISGG---ELYIDKIE 611

QY 636 FIPITATFEAYDLERAQAEVNALFTNTNPRRLKTDVTDYHIDQVSNLACLSDPECLDE 695
DB 612 IILADATFEASDLERAQAVNALFTSNQIGLTKTDVTDYHIDQVSNLVDCLSDPECLDE 671

QY 696 KRELLEKVKYAKRLSDERNLLODNFTSINKQPDFISTNEQSNTSITHEQSEHCWGSN 755
DB 672 KRELSEKVKYAKRLSDERNLLODNFTSINKQPDFISTNEQSNTSITHEQSEHCWGSN 713

```

```

QY 756 ITIQGNDFVKNYVTLPGTFNECYPTLYYQKIGESSELKAYTRYQIRGYIEDSQDLEIYL 815
DB 714 ITIQGDDVFKENYVTLPGTVDECYPTLYYQKIDESKLKAYTRYELRGYIEDSQDLEIYL 773

QY 816 IRYNAKHETLDVPGTESVWPLSVESPIGRGCPNRCAPHFENNPDLDDCCRDGCKCAHHS 875
DB 774 IRYNAKHEIVNVPGTGLMPLSAQSPIGKCGPNRCAPHLEWNPDLDDCCRDGCKCAHHS 833

QY 876 HHFSLDIDIGCTDLHENLGVVWFVKIKTOGHARLGNLFEBEKLPLIGALSVRKAEKK 935
DB 834 HHFTLIDIVGCTDLNEDLGVWIFIKITODGHARLGNLFEBEKLPLIGALSVRKAEKK 893

QY 936 WRDKREKLOLETKRVVTEAKEAVDALFVDSQVNRLOADTNIGMIHAADKLVHRIRAYLS 995
DB 894 WRDKREKLOLETKRVVTEAKEAVDALFVDSQVNRLOADTNIGMIHAADKLVHRIRAYLS 953

QY 996 ELSVIPGVNAEIPFEELEGRITAIISLYDARNVNVKNGDENGLACWNVKGVHDV-QQSHHR 1054
DB 954 ELSVIPGVNAEIPFEELEGRITAIISLYDARNVNVKNGDENGLACWNVKGVHDV-QQSHHR 1013

QY 1055 SVLVIPEWEAEVSQAVRVCPGRGYILRVTAIKYEGYEGCVTTHIEINNTDELKFKNCEE 1114
DB 1014 SVLVIPEWEAEVSQAVRVCPGRGYILRVTAIKYEGYEGCVTTHIEINNTDELKFKNCEE 1073

QY 1115 EYVPTDTGTCNDYTA----HQGTAVCNRSNAGYEDAYEVDVTTASVNYKPTYBEETDVR 1170
DB 1074 EYVPTDTGTCNDYTA----HQGTAVCNRSNAGYEDAYEVDVTTASVNYKPTYBEETDVR 1131

QY 1171 RDNHCEYDRGVNYPPLPAGYMTKELEYEPETDKWIEIGETGKFTIVDSVELLMEE 1228
DB 1132 RENPCESNRGYGDTPLPAGYVTKDLEYFPETDKWIEIGETGKFTIVDSVELLMEE 1189

RESULT 14
US-10-782-020-7
; Sequence 7, Application US/10782020
; Publication No. US20040197916A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-004, A Delta-Endotoxin Gene and
; FILE REFERENCE: 045600/274139
; CURRENT APPLICATION NUMBER: US/10/782,020
; CURRENT FILING DATE: 2004-02-19
; PRIOR APPLICATION NUMBER: 60/448,810
; PRIOR FILING DATE: 2003-02-20
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-782-020-7

Query Match          54.0%; Score 3495.5; DB 16; Length 1189;
Best Local Similarity 56.9%; Pred. No. 6.9e-259;
Matches 716; Conservative 143; Mismatches 290; Indels 109; Gaps 20;

QY 7 NENEII--NALSIAPVNSHSTQMDLSPDARIEDSLCTIAEGNNINPLVSASTVQTGINIAG 64
DB 5 NQNCIPYNCLS-----NPEVLLDGERISTGN-----SSIDISLSLVQ 43

QY 65 RILGLVGPVAGQIASFYSFLVGLMWRGRDQWEIFLEHVEQLINOQITENARNTALARL 124
DB 44 FLVSNF-VPGGFLVGLIDFWGIVGP---SQWDAFLVQIEQLINERIAEFARNAAIL 99

QY 125 QGLGDSFRAYQOQSLDMLNRRDDARTSVLYTQVIALELDFLNAFLPFAIRNOEVLMMV 184
DB 100 EGLGNPNIVYEAKEWEDPNPATRVIDRFRILDGLLDRDIPSDISGPEVPLLSV 159

```


Db 612 IILADATFEASDRLERAQKAVNALFTSSNQIGLKTVDYDHYDQVSNLVDCLSDFECLDE 671
 Qy 696 KRELLEKVKYAKRLSDERLILQDNPFTSINKQPOFISTNEQSNFTSHEQSEHGHWGSEN 755
 Db 672 KRELSEKVKYAKRLSDERLILQDNPFRGINQPD-----RWRGSTD 713
 Qy 756 ITIQGNDVFKENYVTLPGTNECYPTLYOKIGESLKYATRYOLRGYIEDSODLEIYL 815
 Db 714 ITIQGDDVFKENYVTLPGTDECYPTLYOKIDESKLKYATRYELRGYIEDSODLEIYL 773
 Qy 816 IRYNAKHETLDVPGTESVWPLSVESPIGRCEPRNRCAPHFENWPDLDSCDRDGEKCAHS 875
 Db 774 IRYNAKHETLDVPGTESVWPLSVESPIGRCEPRNRCAPHFENWPDLDSCDRDGEKCAHS 833
 Qy 876 HFSLDIDIGCTDLHENLGVWVFKITQOEGHARLGNLEFTEERPLLGCEALSRYKRAEKK 935
 Db 834 HFTFLDIDVGCTDLNEDLGWVWFKITQDGHARLGNLEFTEERPLLGCEALSRYKRAEKK 893
 Qy 936 WRDKREKLQLETKRYVTEAKEAVDALFVDSQYRNLQADTNIGMHAADKLVRHIREAYLS 995
 Db 894 WRDKREKLQLETKRYVTEAKEAVDALFVDSQYRNLQADTNIGMHAADKLVRHIREAYLS 953
 Qy 996 ELSVTPGVNABIPBELEGRIITLSDYDARNVXNGDPNNGIACWNVKGVHDV-QQSHHR 1054
 Db 954 ELSVTPGVNABIPBELEGRIITLSDYDARNVXNGDPNNGIACWNVKGVHDV-QQSHHR 1013
 Qy 1055 SVLVIPBEAARVSOAVRCPGRGVLRTVAYKEGVGSGCVTHIENNTDELKFNCEEE 1114
 Db 1014 SVLVIPBEAARVSOAVRCPGRGVLRTVAYKEGVGSGCVTHIENNTDELKFNCEEE 1073
 Qy 1115 EYVPTDGTGNDYTA----HQTAVCNRSNAGYEDAYEDVDTASVNYKPYEEETDVR 1170
 Db 1074 EYVPTDGTGNDYTA----HQTAVCNRSNAGYEDAYEDVDTASVNYKPYEEETDVR 1131
 Qy 1171 RDNHCEVDRGVNYPPLPAGWMTKELEYFPBTDKWIIEIGTEGKPIVDSVELLMEE 1228
 Db 1132 RENPCESNRGYDGTPLPAGVYTKDLEYFPBTDKWIIEIGTEGKPIVDSVELLMEE 1189

RESULT 10

US-09-972-175-2
 ; Sequence 2, Application US/09972175
 ; Publication No. US20030101482A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baum, James A.
 ; Gilmer, Amy Jelen
 ; Mettus, Anne-Marie Light
 ; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING
 ; LEPIDOPTERAN-ACTIVE-DELTA-ENDOTOXINS
 ; NUMBER OF SEQUENCES: 76
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Arnold, White & Durkee
 ; STREET: P.O. Box 4433
 ; CITY: Houston
 ; STATE: Texas
 ; COUNTRY: USA
 ; ZIP: 77210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/972,175
 ; FILING DATE: 05-Oct-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/337,635
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kitchell, Barbara S.
 ; REGISTRATION NUMBER: 33,928
 ; REFERENCE/DOCKET NUMBER: MSCO:206

TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 512/418-3000
 ; TELEFAX: 512/474-7577
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1189 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 ; US-09-972-175-2
 ; Query Match 54.0%; Score 3500.5; DB 10; Length 1189;
 ; Best Local Similarity 57.0%; Pred. No. 2.8e-259;
 ; Matches 717; Conservative 143; Mismatches 289; Indels 109; Gaps 20;
 ; Qy 7 NENEII--NALSTPAVSNHSTQMDLSPDARIEDSLCTAEGNNINPLVSASTVQTGINIAG 64
 ; Db 5 NQNCIPYNCLS-----NPEVLLDGERISTGN-----SSIDISLSLVQ 43
 ; Qy 65 RILGLVGPAGQIASFYSLVGLWPRGRDQWEI FLEHVEQLINQOITENARNTALRL 124
 ; Db 44 FLVSNF-VPGGGLVGLIDFVWGIVGP---SQMDAFLVQIEQLINERIAEFARNAAL 99
 ; Qy 125 QGLGDSFRAYQOQSLDLEWLENRDARTSVLYTQVIALELDFLNA MPLFAIRNQEVPLIMV 184
 ; Db 100 EGLGNFNIIYVEAFKEWEDPNPATRTVIDRFRILDGLLERDIPSAISGFEVPLSV 159
 ; Qy 185 YQAANLHLLLRDASLFGSEFGLTSQEIORYERYEQVQTRDYSDYCVWEYNTGLNSLRG 244
 ; Db 160 YQAANLHLLLRDASLFGSEFGLTSQEIORYERYEQVQTRDYSDYCVWEYNTGLNSLRG 219
 ; Qy 245 TNAASVRYNQFQRLDLGLVLDLVALPPSYDTYTPINTSAQLTRVYTTDAIGATGVNMA 304
 ; Db 220 STYQDMITVYLRRLDRLTLVLDIAAFPYNDYRNPYIOPVQQLTRVYTTDPL----INFN 275
 ; Qy 305 SMNMYNNAPSFSAIETAVIRSPHLLDFLEQLTIFSTSRWSATRHMTYWRGHTIOSRPI 364
 ; Db 276 POLQSAQLPTFNWMESSAIRNPHLFDILNLTFTD---WFSVGRNFYWGHRVSSLI 332
 ; Qy 365 GGLNLTSTHGSTWTSINPVRLSPFRSDVYVWTESYAGVLL-----WGIYLEPIHGVPTRFN 420
 ; Db 333 GCGNITSPIYCREANQEPFRSFTFNGFVFTLSNPTLRLLQQPWPAPPFPNLRGVEGVFS 392
 ; Qy 421 FRNPQTFE---RGTYNSQPYESPGIQLKXSETLPPETTERPNYESYSHRLSHIGLIS 477
 ; Db 393 --TPTNSFTYRGRTV-----DSUTELPFDNSVPPREGYSHRLCHATFVQ 436
 ; Qy 478 QSRVHV-----PVYSWTHRSADRTNTISSDITQIPLVKSFNLNSGTSVSGPGFTGGDI 533
 ; Db 437 RSGTPTLTGTVFVSWTHRSATLNTIDPERINQIPLVKGFVWGGTSVITGPGFTGGDIL 496
 ; Qy 534 RTNVGSLVSMGLNFNTSLQRYRVRYAASQ-----TMVLRTVVGSGTTFDQGFPS 586
 ; Db 497 RRNTFGDFVSLQVNIINSPIQRYRLRFRVASSRDARVILVTGAASTGVGGQVSNMPLQK 556
 ; Qy 587 TMSANESLTSQSPFAEFP-----VGSASGSQTAGISINNACRGTFPHDKIE 635
 ; Db 557 TMEIGENLTRYTRYTDFSNPFSFRANPDITIGISEQPLFGAG--SISG-----ELYIDKIE 611
 ; Qy 636 FIPIITFAEYDLERAQEAVALFTNTNPRRLKTDVTDYHIDQVSNLVAQLSDEFCLDE 695
 ; Db 612 IILADATFEASDRLERAQKAVNALFTSSNQIGLKTVDYDHYDQVSNLVDCLSDFECLDE 671
 ; Qy 696 KRELLEKVKYAKRLSDERLILQDNPFTSINKQPOFISTNEQSNFTSHEQSEHGHWGSEN 755
 ; Db 672 KRELSEKVKYAKRLSDERLILQDNPFRGINQPD-----RWRGSTD 713
 ; Qy 756 ITIQGNDVFKENYVTLPGTNECYPTLYOKIGESLKYATRYOLRGYIEDSODLEIYL 815
 ; Db 714 ITIQGDDVFKENYVTLPGTDECYPTLYOKIDESKLKYATRYELRGYIEDSODLEIYL 773
 ; Qy 816 IRYNAKHETLDVPGTESVWPLSVESPIGRCEPRNRCAPHFENWPDLDSCDRDGEKCAHS 875

```
Db 220 STYQDMITYNRLRRDLTLVLDIAAFPNYDNRYPQPVQQLTREYVTDPL-----INFN 275
Qy 305 SMWYNNNAPSIAETAVIRSPHLLDFLEQLTIFSTSSRWASATRMWYWRGHTIQSRPI 364
Db 276 PQLOQVAQLPTFNWESSAIRPHLDLNNLTFTD--WFSGRNFYWGGRHVSSLI 332
Qy 365 GGLNTSTHGSTNTSINPVLSPFSRDVYWTESYAGVLL-----WGIYLEPIHGVTVRFN 420
Db 333 GCGNITSPIYGREANQEPFRSFTFNGVFRFTLSNPTLRLLQOPWAPPNLRGVEGVEFS 392
Qy 421 FRNPQTFE---RGTYANSQPYESPGQLQKDSSTELPPTETTERPNYESYSHRLSHIGLIS 477
Db 393 --TPTNSTYVRGTV-----DSLTELPPEDNSVPPREGYSHRLLCHATFVQ 436
Qy 478 QSRVHV-----PVYSWTHRSADRTNTISSDITQIPLVKFNLNSGTSVVSQPGFTGDI 533
Db 437 RSGTPTLTGTVFVSWTHRSATLNTIDPERINQIPLVKGRVNGGTSVIITGPGFTGDI 496
Qy 534 RTNVNGSVLSMGLNFNNTSLQRYRVRVRYAASQ-----TMVLRVTVGGSSTTFDQGFPS 586
Db 497 RRTNFGDFVSLQVNSPITQRYRLRFRYASSRDARVILVTGAASTGVGGQSVNNPQK 556
Qy 587 TMSANESLTSQSFRFAEFP-----VGISASGSQTAGISISNNAGROTFFHDKIE 635
Db 557 TMEIGENLTSRTFRYTDFSNPFSSFRANPDIIIGISEQLPFCAG-SISGG-----ELYDKIE 611
Qy 636 FIPITATFEAYDLERAQAVNALFTNTNPRRLKTDVTDYHIDQVSNLVACLSDFECLDE 695
Db 612 ILLADATFEASDLERAQAVNALFTSSNOIGLTKTDVTDYHIDQVSNLVACLSDFECLDE 671
Qy 696 KRELLEKVKYAKLSDBERNLLQDPNFTSINKQDPFISTNEQSNFTSIHEQSEHGHWGSEN 755
Db 672 KRELSEKVKHAKLSDBERNLLQDPNFRGINRQPD-----RGRWGSTD 713
Qy 756 ITIOEGNDVPEKENYVTLPGTFNFCYPTLYOKIYGESELKAYTRYQLRGYIEDSOILEYL 815
Db 714 ITIOGGDDVPEKENYVTLPGTVDECPYLYOKIDESKLYKAYTRYELRGYIEDSOILEYL 773
Qy 816 IRYNAXHETLDVPGTSEVMPLESPIRCGCEPNRCAPHPEWNPDLDCSCRCDEKCAHNS 875
Db 774 IRYNAXHEIVNVPGTSLWPLSAQSPGKCGCEPNRCAPHLEWNPDLDCSCRCDEKCAHNS 833
Qy 876 HIFSLDIDIGCTDLHENLVVWVFKIKTOGHARLGNLEFIEBKPLLGEALSRVKAEEK 935
Db 834 HHFTLDDVCGCTDLNEDLGWVIFKIKTOGHARLGNLEFIEBKPLLGEALSRVKAEEK 893
Qy 936 WRDKREKLOLETKRYVTEAKEADVALFVDSQVNRQLQADTNIGMTHAADKLVRHIREAYLS 995
Db 894 WRDKREKLOLETNIVYKEAKESVDALFVNSQYDRLOQVDTNIAHAAKKRVHIREAYLP 953
Qy 996 ELSVIPGVNAEIPFEELEGRIITAIISLYDARNVVKNGDFNNGLACWNVKGHDVY-QQSHR 1054
Db 954 ELSVIPGVNAEIPFEELEGRIITAIISLYDARNVVKNGDFNNGLACWNVKGHDVYQSHR 1013
Qy 1055 SVLVIPEWAEVQAVRVCGRGYILRVATYKSGYEGCVTIHEIENNTDELKFKNCBEE 1114
Db 1014 SVLVIPEWAEVQAVRVCGRGYILRVATYKSGYEGCVTIHEIENNTDELKFKNCBEE 1073
Qy 1115 EYVPTDGTGNDYTA-----HQTAVCNRSRAGVEDAYEDVTTASVNYKPYEETVTDYR 1170
Db 1074 EYVNNVTVCNNTGTQEEYEGT--YTSRNOGYDEAYGNNPVPADYASVYEESKSYTDGR 1131
Qy 1171 RDNHCEYDRGVNYPPLPAGMYTKLEYFPETDKVWIEIGETGKFIIVDSVELLMEE 1228
Db 1132 RENPCSNRGYGYTPLPAGVTVYKOLEYFPETDKVWIEIGETGTFIVDSVELLMEE 1189
```

RESULT 9

US-10-200-522-59

; Sequence 59, Application US/10200522

; Publication No. US20030195336A1

; GENERAL INFORMATION:

; APPLICANT: Baum, James A.

```
; APPLICANT: Gilmer, Amy Jelen
; APPLICANT: Mettuss, Anne Marie Light
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE COMPOSITIONS ENCODING LEPIDOPTERAN-TC
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: MECO:213 (11792.0213 DVUS01)
; CURRENT APPLICATION NUMBER: US/10/200,522
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/337,280
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 08/980,071
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 08/757,536
; PRIOR FILING DATE: 1996-11-27
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 59
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant Delta Endotoxin
; US-10-200-522-59
```

Query Match 54.1%; Score 3502.5; DB 14; Length 1189;
Best Local Similarity 57.0%; Pred. No. 2e-259;
Matches 717; Conservative 143; Mismatches 289; Indels 109; Gaps 20;

```
Qy 7 NENEII--NALSIAPVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGINIAG 64
Db 5 NQNCQIPYNCLS-----NPEVLLDGERISTGN-----SSIDISLSLVQ 43
Qy 65 RILGLVGVFPAGQIASFYSLVGLMWRGRDQWEIFLEHVQLINQOITENARNTALARL 124
Db 44 FLVSNE-VFGGGFLVGLIDFVWGIQVGP---SOWDAFLVQIEQLINERIAEFARNAIAANL 99
Qy 125 QGLGDSFRAYQOSLEWLENRRDDARTSVLYTOYIALELDFLNAMPFLFAIRNOEVPLLMV 184
Db 100 EGLGNFNFIYVEAFKEWEEDPNPATRTRVIDRFRILDGLLREDIPISFAISGFEVPLLSV 159
Qy 185 YQAANLHLLLRDASLFGSEGLTSGEIQRYVEROVEQTRDYSYCVENYNTGLNSLRG 244
Db 160 YQAANLHLLLRDASLFGSEGLTSGEIQRYVEROVEQTRDYSYCVENYNTGLNSLRG 219
Qy 245 TNAASVRYNQFRRLDITGLVLDVALPSPYDTRTPINTSAQLTREYVTDATGATGNMA 304
Db 220 STYQDMITYNRLRRDLTLVLDIAAFPNYDNRYPQPVQQLTREYVTDPL-----INFN 275
Qy 305 SMWYNNNAPSIAETAVIRSPHLLDFLEQLTIFSTSSRWASATRMWYWRGHTIQSRPI 364
Db 276 PQLOQVAQLPTFNWESSAIRPHLDLNNLTFTD--WFSGRNFYWGGRHVSSLI 332
Qy 365 GGLNTSTHGSTNTSINPVLSPFSRDVYWTESYAGVLL-----WGIYLEPIHGVTVRFN 420
Db 333 GCGNITSPIYGREANQEPFRSFTFNGVFRFTLSNPTLRLLQOPWAPPNLRGVEGVEFS 392
Qy 421 FRNPQTFE---RGTYANSQPYESPGQLQKDSSTELPPTETTERPNYESYSHRLSHIGLIS 477
Db 393 --TPTNSTYVRGTV-----DSLTELPPEDNSVPPREGYSHRLLCHATFVQ 436
Qy 478 QSRVHV-----PVYSWTHRSADRTNTISSDITQIPLVKFNLNSGTSVVSQPGFTGDI 533
Db 437 RSGTPTLTGTVFVSWTHRSATLNTIDPERINQIPLVKGRVNGGTSVIITGPGFTGDI 496
Qy 534 RTNVNGSVLSMGLNFNNTSLQRYRVRVRYAASQ-----TMVLRVTVGGSSTTFDQGFPS 586
Db 497 RRTNFGDFVSLQVNSPITQRYRLRFRYASSRDARVILVTGAASTGVGGQSVNNPQK 556
Qy 587 TMSANESLTSQSFRFAEFP-----VGISASGSQTAGISISNNAGROTFFHDKIE 635
Db 557 TMEIGENLTSRTFRYTDFSNPFSSFRANPDIIIGISEQLPFCAG-SISGG-----ELYDKIE 611
Qy 636 FIPITATFEAYDLERAQAVNALFTNTNPRRLKTDVTDYHIDQVSNLVACLSDFECLDE 695
```

Qy 121 LARLOGLGDSFRAVQOQSLDLEWLNRRDARTSRVLYTQVIALELDFLNAMPFAIRNOEVP 180
 Db 121 LARLOGLGNSFRAVQOQSLDLEWLNRRDARTSRVLYTQVIALELDFLNAMPFAIRNOEVP 180
 Qy 181 LLMVYAQAANLHLLLDASLFGSEGLTQSEIQRYRQVQEQTRDYSDYCVHYNTGLN 240
 Db 181 LLMVYAQAANLHLLLDASLFGSEGLTQSEIQRYRQVQEQTRDYSDYCVHYNTGLN 240
 Qy 241 SLRGTAASVRYNQFRDLTLGVLDLVALPSPVDTTRYPINTSAQLTRVYTDGATG 300
 Db 241 NLRGTAESWLRYNQFRDLTLGVLDLVALPSPVDTTRYPINTSAQLTRVYTDGATG 300
 Qy 301 V--NMASWYNNAPSAFSAIETAVIRSPHLLDLEQLTIFSTSRWSATRHMYTWRGHT 358
 Db 301 APSGFASWYNNAPSAFSAIETAVIRSPHLLDLEQLTIFSTSRWSATRHMYTWRGHT 358
 Qy 359 IORPPIGGLNTSTHGNTNTSINPRLSPFRSDYVWTVESVAGVLLWGLYLEPHGVPTVR 418
 Db 361 LESTIRGSLSTHGNTNTSINPRLSPFRSDYVWTVESVAGVLLWGLYLEPHGVPTVR 418
 Qy 419 FNRPNQPTFERGTANYSPQVSPCLQKQSETELPEPTEPNERPNYESYSHRLSHIGLISQ 478
 Db 419 FNRPNPLNSL-RGSLTYTIGTGVLQFQSETELPEPTEPNERPNYESYSHRLSNIRLISG 477
 Qy 479 SRVHPVYSWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVVGPGFTGGDIIRTNVN 538
 Db 478 NTLRAPVYSWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVVGPGFTGGDIIRTNVN 537
 Qy 539 GSVLSMGLNFNNLSLQRYRVRYAAQTMVLRVTVGSGTTFDQGFSTMSANESLTSQS 598
 Db 538 GSVLSMGLNFNNLSLQRYRVRYAAQTMVLRVTVGSGTTFDQGFSTMSANESLTSQS 597
 Qy 599 FRPAEPFVGLSAGSQTAGISISNAGQTFHPDKIEFIPITATFEAEYDLERAQEAUNA 658
 Db 598 FRPAEPFVGLSAGSQTAGISISNAGQTFHPDKIEFIPITATFEAEYDLERAQEAUNA 657
 Qy 659 LFTNTPRRLKTDVTDYHIDQVSNLVAQLSDEFCLDEKRELLKVKYAKLSDBERNLLQD 718
 Db 658 LFTSSNQIGLKTVDTHIDRVSNLVECLSDEFCLDEKRELLKVKYAKLSDBERNLLQD 717
 Qy 719 PNFTSINKQDPFTISNQSNFTSISHEQSEHGWSGSENIITQOENVDVFNKENVTLPGTFNE 778
 Db 718 PNFRGINRQLD-----RGWRGSTDTITQGGDDVFNKENVTLPGTFDE 759
 Qy 779 CYPYLYOKIGESLKYTRYQYRGYLEDSDLEIYLIRYNAXHETLDLVPGETSVWPLSV 838
 Db 760 CYPYLYOKIDESLKYTRYQYRGYLEDSDLEIYLIRYNAXHETVYVPGTSLWPLSA 819
 Qy 839 ESPITGRGCEPNRCAPHPEWNPDLDCSCRDGKCAHSHHPSLSDIDIGCTDLHENLGVWV 898
 Db 820 PSPIG-----KCAHSHHPSLSDIDIGCTDLHENLGVWV 853
 Qy 899 FKIKTOGHARLGNLEPIEKPLLGEALSRVRAEKWRDKREKLOLETKRVYTEAKEAV 958
 Db 854 FKIKTOGHARLGNLEPIEKPLLGEALSRVRAEKWRDKREKLEWETINIVYKEAKESV 913
 Qy 959 DALFVDSQYRNLQADTNIGMHAADKLAVHRIEAYLSELSVIPGVNAEIEELEGRIITA 1018
 Db 914 DALFVNSQYRNLQADTNIGMHAADKLAVHRIEAYLSELSVIPGVNAEIEELEGRIFTA 973
 Qy 1019 ISLYDARNVKNKGFNNGLACWNVKGVHDV-QQSHRSVLVPIPEWAEVSVQAVRCPGRG 1077
 Db 974 FSLYDARNVKNKGFNNGLACWNVKGVHDV-QQSHRSVLVPIPEWAEVSVQAVRCPGRG 1033
 Qy 1078 YILRVTAKEGYGEGCVTIHEIENNTDELKPKNCEEEVYPTGTCDNYTA----HQGT 1133
 Db 1034 YILRVTAKEGYGEGCVTIHEIENNTDELKPKNCEEEVYPTGTCDNYTAQTOEEYEGT 1093
 Qy 1134 AVCNSRNAGYEDAYEDVTASVKNPYEETVTDVDRDNHCEYDRGVNYPPLPAGYMT 1193
 Db 1094 --YTSNRNGDYAESNNSVPADYASAYEBKAYTDGRRDNPCESNRGSGDYTLPAGYVT 1151

Qy 1194 KELEYFPETDKWIEIGETGKGFIVDSVELLMEE 1228
 Db 1152 KELEYFPETDKWIEIGETGKGFIVDSVELLMEE 1186
 RESULT 8
 US-09-972-175-59
 ; Sequence 59, Application US/09972175
 ; Publication No. US20030101482A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baum, James A.
 ; Gilmer, Amy Jelen
 ; Mettus, Anne-Marie Light
 ; TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING
 ; LEPIDOPTERAN-ACTIVE-DELTA-ENDOTOXINS
 ; NUMBER OF SEQUENCES: 76
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Arnold, White & Durkee
 ; STREET: P.O. Box 4433
 ; CITY: Houston
 ; STATE: Texas
 ; COUNTRY: USA
 ; ZIP: 77210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/972,175
 ; FILING DATE: 05-Oct-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/337,635
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kitchell, Barbara S.
 ; REGISTRATION NUMBER: 33,928
 ; REFERENCE/DOCKET NUMBER: MECO:206
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 512/418-3000
 ; TELEFAX: 512/474-7577
 ; INFORMATION FOR SEQ ID NO: 59:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1189 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 59:
 US-09-972-175-59
 Query Match 54.1%; Score 3502.5; DB 10; Length 1189;
 Best Local Similarity 57.0%; Pred. No. 2e-259;
 Matches 717; Conservative 143; Mismatches 289; Indels 109; Gaps 20;
 Qy 7 NENEIIT--NALSIPTAVSNHSTOMDLPDARIESLCTIAEGNNINPLVASTVQTGINIAG 64
 Db 5 NQNCIPYNCLS-----NPEEVLDDGERISTGN-----SSIDISLSLVQ 43
 Qy 65 RIILGVIGVPAGQIASFYSFLVGLWPRGRDQWEI FLEHVQELINQOITENARNVALARL 124
 Db 44 FLVSNF-VPGGGFLVGLIDFVWGIVGP--SOWDAFLVQIEQLINERIAEFARNAAILN 99
 Qy 125 QGLGDSFRAVQOQSLDLEWLNRRDARTSRVLYTQVIALELDFLNAMPFAIRNOEVP 184
 Db 100 EGIENNFIYVEAFKWEEDFPNPNATRTVRIDFRILDDGLLERDIPSFALSGFEVPLSV 159
 Qy 185 YAAANLHLLLDASLFGSEGLTQSEIQRYRQVQEQTRDYSDYCVHYNTGLN 244
 Db 160 YAAANLHLLLDASLFGSEGLTQSEIQRYRQVQEQTRDYSDYCVHYNTGLN 219
 Qy 245 YAAANLHLLLDASLFGSEGLTQSEIQRYRQVQEQTRDYSDYCVHYNTGLN 304

APPLICATION NUMBER: US 07/772,027
FILING DATE: 04-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-18051
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
LENGTH: 1207 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-988-462-7

Query Match 87.4%; Score 5659.5; DB 10; Length 1207;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 1079; Conservative 34; Mismatches 89; Indels 5; Gaps 2;
QY 27 MDLSPDARIEDSLCIARGNNINPLVASTVQTGINIAGRILGVLPFAGQIASFYSLV 86
DB 1 MDLDPDARIEDSLCIARGNNIDFVASTVQTGINIAGRILGVLPFAGQIASFYSLV 60
QY 87 GELWPRGRDQWEIFLEHVEQLINQIITENARNTALARLQGLGDSFRAYQOSLEDWLENRD 146
DB 61 GELWPRGRDQWEIFLEHVEQLINQIITENARNTALARLQGLGDSFRAYQOSLEDWLENRD 120
QY 147 DARTSRVLYTQYIALELDFLNAPLFAIRNQEVPLLMVYAQAANLHLLLRDASLFGSEF 206
DB 121 DARTSRVLYTQYIALELDFLNAPLFAIRNQEVPLLMVYAQAANLHLLLRDASLFGSEF 180
QY 207 GLTSQEIQRVYRQVEOTRYSYCVENYNTGNSLGTNAASWVRVYNQFRRDLTLGLVD 266
DB 181 GLTSQEIQRVYRQVEOTRYSYCVENYNTGNSLGTNAASWVRVYNQFRRDLTLGLVD 240
QY 267 LVLPFSYDTRTPINTSAQLTREVTYDAIGATGVNMAWNNNAPSAIETAVIRS 326
DB 241 LVLPFSYDTRTPINTSAQLTREVTYDAIGATGVNMAWNNNAPSAIETAAIRS 300
QY 327 PHLLDFLEQLTIFSTSRWSATRMVTYWRGHTTQSRPIGGGLNTSTHGSNTNINPVRLS 386
DB 301 PHLLDFLEQLTIFSTSRWSATRMVTYWRGHTTQSRPIGGGLNTSTHGSNTNINPVRLR 360
QY 387 FFSRDVYRTESYAGVLWGILYLPFIHQVPTVRNFRNPONTFPGTANYSQPYESGLQL 446
DB 361 FASRDVYRTESYAGVLWGILYLPFIHQVPTVRNFRNPONTFPGTANYSQPYESGLQL 420
QY 447 KDSETELPPTETRPNYESYSHRLSHGLISQSRVHPVYVSWTHRSADRNTNTSSDSITQ 506
DB 421 KDSETELPPTETRPNYESYSHRLSHGLISQSRVHPVYVSWTHRSADRNTNTGPNRITQ 480
QY 507 IPLVKSFLNLSGTSVWSGPGTGGDIIRTNVNGSVLSMGLNFNNISLQRYVRVRYAASQ 566
DB 481 IPWVKASELPQGTIVVRGPGFTGGDILRRNTTGGFGPIRVTVNGPLTQRYRIGFRYASTV 540
QY 567 TMLRVTVGSGTTFDQGFPTMSANESLTSQSRFPAEPFVIGISASGQ-TAGISISNAG 625
DB 541 DFDFFVSRGTTVNNFRFLTMSGDELKYNFVRRAFTPTFTQTQIDIRTSIQGLSG 600
QY 626 RQTFHFQKIEFIPITATFEAYDLERAQEAVALFTNTNPRRLKTDVTDYHIDQVSNLVA 685
DB 601 NGEVYIDKIEIIPATFEAYDLERAQEAVALFTNTNPRRLKTDVTDYHIDQVSNLVA 660
QY 686 CLSDEFCLDEKRELLEKVKYAKRLSDERNLLQDPNFTSINKQPDFISTNQSNFTSIHQ 745
DB 661 CLSDEFCLDEKRELLEKVKYAKRLSDERNLLQDPNFTSINKQPDFISTNQSNFTSIHQ 720
QY 746 SEHGWSGSENIITQEGNDVFNKENVYVTLPGTFNECYPTLYLQKIGESLKAYTRYQLRGYI 805
DB 721 SEHGWSGSENIITQEGNDVFNKENVYVTLPGTFNECYPTLYLQKIGESLKAYTRYQLRGYI 780

QY 806 EDSQDLIELYLRYNKAKHETLDVPGTESVMPLSVESPIGRGCEPNRCAPHPWNPDLDCSC 865
DB 781 EDSQDLIELYLRYNKAKHETLDVPGTESVMPLSVESPIGRGCEPNRCAPHPWNPDLDCSC 840
QY 866 RDEKCAHSHHPSLDDIGCTDLHENLGVVVFVKIKTQEGHARLGNLEFIEKPLLGEA 925
DB 841 RDEKCAHSHHPSLDDIGCTDLHENLGVVVFVKIKTQEGHARLGNLEFIEKPLLGEA 900
QY 926 LSRVRAEKKWRDKREKLOLETKRVVTEAKEAVALFVDSQVNRQLQADTNIGMHAADKL 985
DB 901 LSRVRAEKKWRDKREKLOLETKRVVTEAKEAVALFVDSQVNRQLQADTNIGMHAADKL 960
QY 986 VHRIRAYLSSELVPIGVNAEIPFEELEGRITITAIISYDARNVVKNGDFNGLACWNVKGH 1045
DB 961 VHRIRAYLSSELVPIGVNAEIPFEELEGRITITAIISYDARNVVKNGDFNGLACWNVKGH 1020
QY 1046 VDVOQSHHPSLVIPWEAEVSOAVRCPGRGVILRVATYKEGCGVGTIHEIENNTDE 1105
DB 1021 VDVOQSHHPSLVIPWEAEVSOAVRCPGRGVILRVATYKEGCGVGTIHEIENNTDE 1080
QY 1106 LKFKNCEEEVYPTDGTCTNDYTAHQGTA----VCNSRNAGYEDAYEVDTTASVNYKPTY 1161
DB 1081 LKFKNCEEEVYPTDGTCTNDYTAHQGTA----VCNSRNAGYEDAYEVDTTASVNYKPTY 1140
QY 1162 EETTYTDVRDNRHCEYDRGVNYPPLPAGYMTKELFYPETDKWIEIGETGKFIVDSV 1221
DB 1141 EETTYTDVRDNRHCEYDRGVNYPPLPAGYMTKELFYPETDKWIEIGETGKFIVDSV 1200
QY 1222 ELLLMEE 1228
DB 1201 ELLLMEE 1207
RESULT 7
US-09-826-660-23
; Sequence 23, Application US/09826660
; Patent No. US20010026940A1
; GENERAL INFORMATION:
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Stelman, Steven J.
; APPLICANT: Narva, Kenneth E.
; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins
; FILE REFERENCE: MA-714XC2D1
; CURRENT APPLICATION NUMBER: US/09/826,660
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/178,252
; PRIOR FILING DATE: 1998-10-23
; PRIOR APPLICATION NUMBER: 60/065,215
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/076,445
; PRIOR FILING DATE: 1998-03-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1186
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Toxin encoded by synthetic B.t. gene
US-09-826-660-23

Query Match 78.8%; Score 5108; DB 9; Length 1186;
Best Local Similarity 80.6%; Pred. No. 0;
Matches 995; Conservative 76; Mismatches 108; Indels 56; Gaps 8;
QY 1 LTNRNKNEIINALSIPAVNSHSTOMDLSPDARIEDSLCIAEGNNINPLVASTVGTGI 60
DB 1 MTNRNKNENIINALSIPAVNSHSAQWNLSTDAIEDSLCIAEGNNIDPFVASTVGTGI 60
QY 61 NIAGRILGVLPFAGQIASFYSLVGEAWPRGRDQWEIFLEHVEQLINQIITENARNTA 120
DB 61 NIAGRILGVLPFAGQIASFYSLVGEAWPRGRDQWEIFLEHVEQLIRQQVNTNRDIA 120

; CURRENT APPLICATION NUMBER: US/10/809,953
 ; CURRENT FILING DATE: 2004-03-26
 ; PRIOR APPLICATION NUMBER: US/09/661,016
 ; PRIOR FILING DATE: 2000-09-13
 ; PRIOR APPLICATION NUMBER: PCT/EP90/00905
 ; PRIOR FILING DATE: 1990-05-30
 ; PRIOR APPLICATION NUMBER: GB 89401499.2
 ; PRIOR FILING DATE: 1989-05-31
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 10
 ; LENGTH: 1228
 ; TYPE: PRT
 ; ORGANISM: Bacillus thuringiensis
 US-10-809-953-10

Query Match 88.68; Score 5742; DB 16; Length 1228;
 Best Local Similarity 89.18; Pred. No. 0;
 Matches 1098; Conservative 35; Mismatches 90; Indels 10; Gaps 3;

QY	1	LTSNRKNEINALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVSASTVGTGI	60
DB	1	LTSNRKNEIN-----AVSNHSAQMDLAPDARIEDSLCIAEGNNIDPFVASTVGTGI	55
QY	61	NIAGRILGLVGPAGQIAGFYSLVGLMPRGDQWEIFLEHVEQLINQOITENARNTA	120
DB	56	NIAGRILGLVGPAGQIAGFYSLVGLMPRGDQWEIFLEHVEQLINQOITENARNTA	115
QY	121	LARLOGLGDSFRAYQOQSLDLENRDDARTSVLTQYIALELDFNAMPLFAIRNOEVP	180
DB	116	LARLOGLGDSFRAYQOQSLDLENRDDARTSVLTQYIALELDFNAMPLFAIRNOEVP	175
QY	181	LLMYYAQAANLHLLLDASLFGSEFGLTSQEIQRVYVEROVEDTSDYSCVEMWNTGLN	240
DB	176	LLMYYAQAANLHLLLDASLFGSEFGLTSQEIQRVYVEROVEDTSDYSCVEMWNTGLN	235
QY	241	SLRGNTAASWVRYNQPRDRLTLGLVDLVALFPSTYDTRTYPINTSAQLTRVYTDAGATG	300
DB	236	SLRGNTAASWVRYNQPRDRLTLGLVDLVALFPSTYDTRTYPINTSAQLTRVYTDAGATG	295
QY	301	VNWSMWNYNNAAPSFAIETAVTRSPHLADLEQLTIFSTSSRSASATRMWYRGHTIQ	360
DB	296	VNWSMWNYNNAAPSFAIETAVTRSPHLADLEQLTIFSTSSRSASATRMWYRGHTIQ	355
QY	361	SRPIGGGLNTSTHGSTNTSINPVRSLFFSRDVTYVWTSYAGVLLWGLIYLEPIHGVPVTFRN	420
DB	356	SRPIGGGLNTSTHGATNTSINPVTLRASRDVVTESYAGVLLWGLIYLEPIHGVPVTFRN	415
QY	421	FRNPQTFERTANYSPQYSPGLQKDSQETELPPETTERPNVYESYSHRSLHIGLSQSR	480
DB	416	FTNPQNSDRGTANYSPQYSPGLQKDSQETELPPETTERPNVYESYSHRSLHIGLSQSR	475
QY	481	VHVPVYSWTHRSADRTNTTSSDSITQIPLVKSFNLSNGTSVSGPGTGGDIIRTNVGS	540
DB	476	VNVPVYSWTHRSADRTNTTIPNRTITQIPMKASGELPQGTITVVRGPGTGGDIIRTNVGS	535
QY	541	VLSMGLNFNTSLQRVYRVYAAQTMVLRVTVGSTTFDQGPSTMSANESILTSQSPR	600
DB	536	FGIRVTVNGPLTQRYAIGFYASTVDFDFVVRGGTGVNPFRLTMSGDELKYGNFV	595
QY	601	FAEFPVGISAGSQ-TAGISINNAGSQTFHFKIEFIPITATFEAYDILERAQEAVAL	659
DB	596	RRAFTPTFTQDIQIIRTSIQGLSGNGEVYIDKIEIIPVATFEAYDILERAQEAVAL	655
QY	660	FTYTNPRRLKTDVTDHIDQVSNLVACLDSFCLDEKRELLEKVKYAKRLSDERNLLQDP	719
DB	656	FTYTNPRRLKTDVTDHIDQVSNLVACLDSFCLDEKRELLEKVKYAKRLSDERNLLQDP	715
QY	720	NFTSINKQDPFISTNEOSNTFSTHEQSEHGHWGSENITIOEGNDVFKENYVTLPGTFNEC	779
DB	716	NFTSINKQDPFISTNEOSNTFSTHEQSEHGHWGSENITIOEGNDVFKENYVTLPGTFNEC	775
QY	780	YPTYLYQKIGESSELKAVTRYQLRGYIBDSQDLLEYILIRYNKAKHETLDVPGTGSWPLSV	839

DB	776	YPTYLYQKIGESSELKAVTRYQLRGYIBDSQDLLEYILIRYNKAKHETLDVPGTGSWPLSV	835
QY	840	SPIGRCCEPNRCAPHFEPWNPDLDCSCRDGKCAHSHHFLSDIDIGCTDLHENLGVWVVF	899
DB	836	SPIGRCCEPNRCAPHFEPWNPDLDCSCRDGKCAHSHHFLSDIDVCGCTDLHENLGVWVVF	895
QY	900	KIKTQEGHARLGNLEFIEEKPLLEALSRVKRAEKWRDKREKLQLETKRVYTEAKEAVD	959
DB	896	KIKTQEGHARLGNLEFIEEKPLLEALSRVKRAEKWRDKREKLQLETKRVYTEAKEAVD	955
QY	960	ALFVDSQYNRLQADTNIGMHAADKLVRHIREAYLSLSVPGVNABIFEELEGRITAI	1019
DB	956	ALFVDSQYDRLOADTNIGMHAADKLVRHIREAYLSLSVPGVNABIFEELEGRITAI	1015
QY	1020	SLYDARNVKNNGDPNGLACWNVKGVHDVQOQSHRSVLVPEWEAEVSQAVRCPGRGYI	1079
DB	1016	SLYDARNVKNNGDPNGLACWNVKGVHDVQOQSHRSVLVPEWEAEVSQAVRCPGRGYI	1075
QY	1080	LRVTAYKEGYGEGCVTTHIEIENNTDELKPKNCEEEVYPTDTGTCDNYTAHQGTATA	1135
DB	1076	LRVTAYKEGYGEGCVTTHIEIENNTDELKPKNCEEEVYPTDTGTCDNYTAHQGTATA	1135
QY	1136	CNSRNGYEDAYVDVTTASVNYKPTVEETTYTVDRDNHCEYDRGVYNYPLPAGYWK	1195
DB	1136	CNSRNGYEDAYVDVTTASVNYKPTVEETTYTVDRDNHCEYDRGVYNYPLPAGYWK	1195
QY	1196	LEYFPETDKWIEIGETGKFIVDSEVLLMEE	1228
DB	1196	LEYFPETDKWIEIGETGKFIVDSEVLLMEE	1228

RESULT 6
 US-09-988-462-7
 ; Sequence 7, Application US/09988462
 ; Publication No. US20030046726A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kozziel, Michael G.
 ; Desai, Nalini M.
 ; Lewis, Kelly S.
 ; Kramer, Vance C.
 ; Warren, Gregory W.
 ; Evola, Stephen V.
 ; Crossland, Lyle D.
 ; Wright, Martha S.
 ; Merlin, Ellis J.
 ; Launis, Karen L.
 ; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
 ; INSECTICIDAL ACTIVITY IN MAIZE
 ; NUMBER OF SEQUENCES: 94
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Syngenta Biotechnology, Inc.
 ; STREET: 3054 Cornwalis Road
 ; CITY: Research Triangle Park
 ; STATE: NC
 ; COUNTRY: USA
 ; ZIP: 27709
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/988,462
 ; FILING DATE: 20-NOV-2003
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 09/547,422
 ; FILING DATE: 11-APR-2000
 ; APPLICATION NUMBER: US 08/459,504
 ; FILING DATE: 02-JUN-1995
 ; APPLICATION NUMBER: US 07/951,715
 ; FILING DATE: 25-SEP-1992


```
Db 958 DALFVDSQYDLQADTNIGMHAADKLVRHREAYLSLSVPGVNAEIPFEELEGRITTA 1017
Qy 1019 ISLYDARNVKNVGNFNNGLACWNVKGVHDVVOQSHRSVLVIPEWEAEVSQAVRVCGRGY 1078
Db 1018 ISLYDARNVKNVGNFNNGLACWNVKGVHDVVOQSHRSVLVIPEWEAEVSQAVRVCGRGY 1077
Qy 1079 ILRVATYKEGYGEGCVTHIEIENNTDELKFNCEEEVYPTDCTCNDYTAHQGTAVCNS 1138
Db 1078 ILRVATYKEGYGEGCVTHIEIENNTDELKFNCEEEVYPTDCTCNDYTAHQGTAVCNS 1137
Qy 1139 RNAGYDAYEVDVTTASVNYKPTVEEETVTVRRDNHCEYDRGVNYPPLPAGYMTKELEY 1198
Db 1138 RNAGYDAYEVDVTTASVNYKPTVEEETVTVRRDNHCEYDRGVNYPPLPAGYMTKELEY 1197
Qy 1199 FPETDKWIEIGTEGKFI VDSVELLMEE 1228
Db 1198 FPETDKWIEIGTEGKFI VDSVELLMEE 1227

RESULT 4
US-10-926-819-8
; Sequence 8, Application US/10926819
; Publication No. US20050049410A1
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine
; APPLICANT: Hargiss, Tracy
; APPLICANT: Koziel, Michael G.
; APPLICANT: Duck, Nicholas B.
; APPLICANT: Carr, Brian
; TITLE OF INVENTION: AXMI-003, A Delta-Endotoxin Gene and
; TITLE OF INVENTION: Methods for Its Use
; FILE REFERENCE: 045600/281577
; CURRENT APPLICATION NUMBER: US/10/926,819
; CURRENT FILING DATE: 2004-08-26
; PRIOR APPLICATION NUMBER: 60/498,518
; PRIOR FILING DATE: 2003-08-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1228
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis serovar entomocidus
US-10-926-819-8

Query Match 88.78; Score 5745; DB 17; Length 1228;
Best Local Similarity 89.14; Pred. No. 0;
Matches 1098; Conservative 35; Mismatches 90; Indels 10; Gaps 3;

Qy 1 LTSNRKNENIINALSIPAVSNHSTOMDLSPDARIEDSLCIAEGNNINPLVSASTVQTGI 60
Db 1 MTSNRKNENIIN-----AVSNHSAQMDLLPDARIEDSLCIAEGNNIDPFVSASTVQTGI 55
Qy 61 NIAGRLGLVGVFPAGQIASFYSLVGLWPRGRDQWEIFLEHVQELINQOITENARNTA 120
Db 56 NIAGRLGLVGVFPAGQIASFYSLVGLWPRGRDQWEIFLEHVQELINQOITENARNTA 115
Qy 121 LARLOGLDSPRAYQOQSLWLENRRDARTSVLYTQYIALELDPLNAPLFAIRNOEVP 180
Db 116 LARLOGLDSPRAYQOQSLWLENRRDARTSVLYTQYIALELDPLNAPLFAIRNOEVP 175
Qy 181 LMVYAQAANLHLLLDASLFGSEFGLTQEIORYYEROVEOTRDYSDYCVEMWNTGLN 240
Db 176 LMVYAQAANLHLLLDASLFGSEFGLTQEIORYYEROVEOTRDYSDYCVEMWNTGLN 235
Qy 241 SLRGNTAASVRYNQFRDLTLGLVDLVALFPSPYDTRTPINTSAQLTREVYTDALGATG 300
Db 236 SLRGNTAASVRYNQFRDLTLGLVDLVALFPSPYDTRTPINTSAQLTREVYTDALGATG 295
Qy 301 VNMASWNNYNNAPSFSAIETAVIRSRPHLLDFLEQLTIFTSRWSATRMWTRGHITQ 360
Db 296 VNMASWNNYNNAPSFSAIAEAAIRSRPHLLDFLEQLTIFTSRWSATRMWTRGHITQ 355
```

```
Qy 361 SRPIGGGLTNTSHGSTNTSINPRLSFFGRDVTWTSYAGVLLWGIYLEPIHGVPVTRFN 420
Db 356 SRPIGGGLTNTSHGATNTSINPVLTRFASRDVYRTSYAGVLLWGIYLEPIHGVPVTRFN 415
Qy 421 FRNPQNTFPGTANYSQPYESPGLOIKDSETELPPETTERPNYESYSHRLSHLGLISQSR 480
Db 416 FTNPQNISDRGTANYSQPYESPGLOIKDSETELPPETTERPNYESYSHRLSHLGLISQSR 475
Qy 481 VHVVPVSWTHRSADRNTTSSDSITQIPLVKSFNLSGTSVSVSGPGFTGGDIIRTNVNGS 540
Db 476 VNVVPVSWTHRSADRNTTGPNRITQIPMKASELPQGTIVVRGPGFTGGDIIRTNVNGS 535
Qy 541 VLSMGLNFNNTSLQYRVRYAASQTMVLRVTVGSGTTFDQGFPPSTMSANESLTQSFR 600
Db 536 FGBIRVTAVGPLTQVRIGFRYASTVDFDFVSRGGTIVNFRFLRTMNSGDELKYGNFV 595
Qy 601 FAEFPVGISASGSO-TAGISISNNAGROTFFHFKIEFIPITATFEAYDLERAQEAVAL 659
Db 596 RRAFTTPTFTQIQTDIRTSIOGLSGNGEYVYDIKIEIIPVTATFEAYDLERAQEAVAL 655
Qy 660 FTNTNPRRLKTDVTDYHIDQVSNLVACLSDFCDEKRELLEKVKYAKRLSDERNLLQDP 719
Db 656 FTNTNPRRLKTDVTDYHIDQVSNLVACLSDFCDEKRELLEKVKYAKRLSDERNLLQDP 715
Qy 720 NFTSINKQPDFISTNEQSNFTSIHQSEHGWMGSENITIQEGNDVFKENVYVTLPGTFNEC 779
Db 716 NFTSINKQPDFISTNEQSNFTSIHQSEHGWMGSENITIQEGNDVFKENVYVTLPGTFNEC 775
Qy 780 YPTLYQKIGESSELKAYTRYQLRGYIEDSQDLIELIYLRYNKAKHETLDVPGTESVWPLSVE 839
Db 776 YPTLYQKIGESSELKAYTRYQLRGYIEDSQDLIELIYLRYNKAKHETLDVPGTESVWPLSVE 835
Qy 840 SPIGRGCEPNRCAPHEWNPDLDCSCRDGCKCAHSHHFLSDIDIGCTDLHENLGVWVWF 899
Db 836 SPIGRGCEPNRCAPHEWNPDLDCSCRDGCKCAHSHHFLSDIDVGCCTDLHENLGVWVWF 895
Qy 900 KIKTOGHARLGNLEFIEEKPLIGEALSRVKRAEKKWRDKREKLQLETKRVTYTAKEAVD 959
Db 896 KIKTOGHARLGNLEFIEEKPLIGEALSRVKRAEKKWRDKREKLQLETKRVTYTAKEAVD 955
Qy 960 ALFVDSQYDLQADTNIGMHAADKLVRHREAYLSLSVPGVNAEIPFEELEGRITTA 1019
Db 956 ALFVDSQYDLQADTNIGMHAADKLVRHREAYLSLSVPGVNAEIPFEELEGRITTA 1015
Qy 1020 SLVDARNVKNVGNFNNGLACWNVKGVHDVVOQSHRSVLVIPEWEAEVSQAVRVCGRGY 1079
Db 1016 SLVDARNVKNVGNFNNGLACWNVKGVHDVVOQSHRSVLVIPEWEAEVSQAVRVCGRGY 1075
Qy 1080 LRVATYKEGYGEGCVTHIEIENNTDELKFNCEEEVYPTDCTCNDYTAHQGTAVCNS 1135
Db 1076 LRVATYKEGYGEGCVTHIEIENNTDELKFNCEEEVYPTDCTCNDYTAHQGTAVCNS 1135
Qy 1136 CNSRNAGYDAYEVDVTTASVNYKPTVEEETVTVRRDNHCEYDRGVNYPPLPAGYMTKE 1195
Db 1136 CNSRNAGYDAYEVDVTTASVNYKPTVEEETVTVRRDNHCEYDRGVNYPPLPAGYMTKE 1195
Qy 1196 LEYFPETDKWIEIGTEGKFI VDSVELLMEE 1228
Db 1196 LEYFPETDKWIEIGTEGKFI VDSVELLMEE 1228

RESULT 5
US-10-809-953-10
; Sequence 10, Application US/10809953
; Publication No. US20040181825A1
; GENERAL INFORMATION:
; APPLICANT: Van Mellaert, Herman
; APPLICANT: Botterman, Johan
; APPLICANT: Van Rie, Jeroen
; APPLICANT: Joos, Henk
; TITLE OF INVENTION: RECOMBINANT PLANT EXPRESSING NON-COMPETITIVELY BINDING BT INSECTIC
; TITLE OF INVENTION: CRYSTAL PROTEINS
; FILE REFERENCE: 021565-078
```

```
Qy 541 VLSMGLNPNNTSLQRYRVRVRYAASQTMVLRTVVGSGTTFDQGFPTMSANESLTSQSF 600
Db 541 VLSMGLNPNNTSLQRYRVRVRYAASQTMVLRTVVGSGTTFDQGFPTMSANESLTSQSF 600
Qy 601 FAEPFVGISAGSGTAGISINNAAGRTFHFDPKIEFIPITATFAEYDLERAQAVNALF 660
Db 601 FAEPFVGISAGSGTAGISINNAAGRTFHFDPKIEFIPITATFAEYDLERAQAVNALF 660
Qy 661 TINTNPRRLKTDVTDYHIDQVSNLVACLSDEFCLDEKRELLKVKYAKRLSDERNLLQDPN 720
Db 661 TINTNPRRLKTDVTDYHIDQVSNLVACLSDEFCLDEKRELLKVKYAKRLSDERNLLQDPN 720
Qy 721 FTSINKQPDFISTNEQNSFTSIHQSEHGWMGSENITIQEGNDVFKENYVTLPGTFNECY 780
Db 721 FTSINKQPDFISTNEQNSFTSIHQSEHGWMGSENITIQEGNDVFKENYVTLPGTFNECY 780
Qy 781 PTLYLYQKIGSELKAYTRYQLRGYIEDSQDLLEIYLIRYNAKHETLDVPGTESVWPLSVES 840
Db 781 PTLYLYQKIGSELKAYTRYQLRGYIEDSQDLLEIYLIRYNAKHETLDVPGTESVWPLSVES 840
Qy 841 PIGRCGEPNRCAPHFENWPDLDSCRDGEKCAHSHHFSLDIDIGCTDLHENLGVWVVK 900
Db 841 PIGRCGEPNRCAPHFENWPDLDSCRDGEKCAHSHHFSLDIDIGCTDLHENLGVWVVK 900
Qy 901 IKTOEGHARLGNLEFIEBKPLLGALSRVKAERKWRDKREKLETKRVVYTEAKEAVDA 960
Db 901 IKTOEGHARLGNLEFIEBKPLLGALSRVKAERKWRDKREKLETKRVVYTEAKEAVDA 960
Qy *961 LFDVDSQYNRLOADTNIGMHAADKLVRHREAYLSELVPGVNAEIFELEGRITTAIS 1020
Db 961 LFDVDSQYNRLOADTNIGMHAADKLVRHREAYLSELVPGVNAEIFELEGRITTAIS 1020
Qy 1021 LYDARNVVKNGDFNGLACVNVKGVHDVQOSHRSVLVPEWEAEVSOAVRVCPRGVIL 1080
Db 1021 LYDARNVVKNGDFNGLACVNVKGVHDVQOSHRSVLVPEWEAEVSOAVRVCPRGVIL 1080
Qy 1081 RVTAAYKEGYGEGCVTIHEIENNTDELKFKNCEEEVYPTDGTGNDYTAHQGTAAACNSRN 1140
Db 1081 RVTAAYKEGYGEGCVTIHEIENNTDELKFKNCEEEVYPTDGTGNDYTAHQGTAAACNSRN 1140
Qy 1141 AGYEDAYEDVTTASVNVKPYEETTYDVRDNCHEYDRGVNYPPLPAGMYKELEYFP 1200
Db 1141 AGYEDAYEDVTTASVNVKPYEETTYDVRDNCHEYDRGVNYPPLPAGMYKELEYFP 1200
Qy 1201 ETDKRWIEIGETEGKEFIVDSVELLMEE 1228
Db 1201 ETDKRWIEIGETEGKEFIVDSVELLMEE 1228

RESULT 3
US-10-428-961-63
; Sequence 63, Application US/10428961
; Publication No. US2003023711A1
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Rupar, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin
; TITLE OF INVENTION: Polynucleotides, Compositions, and Methods of Use (Amended)
; FILE REFERENCE: MECO201--1
; CURRENT APPLICATION NUMBER: US/10/428,961
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/661,322
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: 60/153,995
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 63
; LENGTH: 1227
```

```
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-428-961-63
```

```
Query Match 91.3%; Score 5912.5; DB 15; Length 1227;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1127; Conservative 36; Mismatches 62; Indels 5; Gaps 3;
```

```
Qy 1 LTSNRKNENIINALSIPAVSNHSTOMDLSPDARIEDSLCIAEGNNINPLVSASTVGTGI 60
Db 1 LTSNRKNENIINALSIPAVSNHSAQWNLSTDAIEDSLCIAEGNNIDPFVSASTVGTGI 60
Qy 61 NIAGRILGVGVPPAGQIASFYSLVAGELWPRGRDQWEIFLEHVEQLINOQITENARNTA 120
Db 61 NIAGRILGVGVPPAGQIASFYSLVAGELWPRGRDQWEIFLEHVEHLIRQOVTEENTDTA 120
Qy 121 LARLOGLGDSFRAYQOSLEDWLENRDDARTSVLYTOYIALELDFLAMPFLAIRNOEVP 180
Db 121 LARLOGLGDSFRAYQOSLEDWLENRDDARTSVLYTOYIALELDFLAMPFLAIRNOEVP 180
Qy 181 LLMVYQAANLHLLLLDASLFGSEFGLTSQEIQRVYVERQVEOTRDYSDYCVWEYNTGLN 240
Db 181 LLMVYQAANLHLLLLDASLFGSEFGLTSQEIQRVYVERQVEKTRYSYDVCARWYNTGLN 240
Qy 241 SLRGTNAAWVRVYNQPRRDLTLGVLDLVALPSPYDTRTYPTINTSAQLTRVYVYDAIGATG 300
Db 241 NLRGTNAESWLRVYNQPRRDLTLGVLDLVALPSPYDTRTYPTINTSAQLTRVYVYDAIGATG 300
Qy 301 V--NMASMWVNNAPSFSAIETAVRSPLHLDFLEQLTIFSTSSRSWATPHMTYWGHT 358
Db 301 APGFASTNFWNNAPSFSAIEAAVIRPPLHDFPQLTIFSVLSRWSNTQYMNWYVGH 360
Qy 359 IQSRPFGGGLNTSTHGSTNTSINPRLSPFRSDVYWTESYAGVLLWGIYLEPIHGVPVTR 418
Db 361 LEERTIGSLSTWHTGNTSINPVLQFTSRDVRYESPAGINI--LLTTPVNGVPEWAR 418
Qy 419 FNFRNPONTFERGTANYQSPYSPGLQKDSATELPETTERPNYSESYSHRLSHIGLISQ 478
Db 419 FNWRNPLNSL-RGSLLYTIGYTGVTGQTFDSETELPETTERPNYSESYSHRLSNIRLISG 477
Qy 479 SRVHVVPVYSWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVVGSGFGTGGDIIRNVN 538
Db 478 NTLRAPVYSWTHRSADRTNTISSDSITQIPLVKSFNLSGTSVVGSGFGTGGDIIRNVN 537
Qy 539 GSVLSMGLNPNNTSLQRYRVRVRYAASQTMVLRTVVGSGTTFDQGFPTMSANESLTSQS 598
Db 538 GSVLSMGLNPNNTSLQRYRVRVRYAASQTMVLRTVVGSGTTFDQGFPTMSANESLTSQS 597
Qy 599 FRPAEPFVGISAGSGTAGISINNAAGRTFHFDPKIEFIPITATFAEYDLERAQAVNA 658
Db 598 FRPAEPFVGISAGSGTAGISINNAAGRTFHFDPKIEFIPITATFAEYDLERAQAVNA 657
Qy 659 LFTNTPRRLKTDVTDYHIDQVSNLVACLSDEFCLDEKRELLKVKYAKRLSDERNLLQD 718
Db 658 LFTNTPRRLKTDVTDYHIDQVSNLVACLSDEFCLDEKRELLKVKYAKRLSDERNLLQD 717
Qy 719 PNFTSINKQPDFISTNEQNSFTSIHQSEHGWMGSENITIQEGNDVFKENYVTLPGTFNE 778
Db 718 PNFTSINKQPDFINSNNEQNSFTSIHQSEHGWMGSENITIQEGNDVFKENYVTLPGTFNE 777
Qy 779 CYPTLYLYQKIGSELKAYTRYQLRGYIEDSQDLLEIYLIRYNAKHETLDVPGTESVWPLSV 838
Db 778 CYPTLYLYQKIGSELKAYTRYQLRGYIEDSQDLLEIYLIRYNAKHETLDVPGTESVWPLSV 837
Qy 839 ESPIGRCGEPNRCAPHFENWPDLDSCRDGEKCAHSHHFSLDIDIGCTDLHENLGVWV 898
Db 838 ESPIGRCGEPNRCAPHFENWPDLDSCRDGEKCAHSHHFSLDIDIGCTDLHENLGVWV 897
Qy 899 FKIKTOEGHARLGNLEFIEBKPLLGALSRVKAERKWRDKREKLETKRVVYTEAKEAV 958
Db 898 FKIKTOEGHARLGNLEFIEBKPLLGALSRVKAERKWRDKREKLETKRVVYTEAKEAV 957
Qy 959 DALFVDSQYNRLOADTNIGMHAADKLVRHREAYLSELVPGVNAEIFELEGRITTA 1018
```

```

Db 61 NIAGRIILGVLPAGFAGFASFLVGLWPRGRDQWEI FLEHVEQLINQOITENARNTA 120
Qy 121 LARLOGLGDSFRAYQOQSLDLENRDDRTRSVLYTOYIALELDFNAMPLFAIRNOEVP 180
Db 121 LARLOGLGDSFRAYQOQSLDLENRDDRTRSVLYTOYIALELDFNAMPLFAIRNOEVP 180
Qy 181 LLMVYQAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVEQTRDYSYCVWEYNTGLN 240
Db 181 LLMVYQAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVEQTRDYSYCVWEYNTGLN 240
Qy 241 SLRGTAASWVRVNOQFRDRLTLGVLDLVALFPSTYTRTPINTSAQLTRREYVYDAIGATG 300
Db 241 SLRGTAASWVRVNOQFRDRLTLGVLDLVALFPSTYTRTPINTSAQLTRREYVYDAIGATG 300
Qy 301 VNMAWMNNYNNAPSFSAIETAVIRSPHLLDLEQLTIFSTSSRWSSATRMHTYWRGHTIQ 360
Db 301 VNMAWMNNYNNAPSFSAIETAVIRSPHLLDLEQLTIFSTSSRWSSATRMHTYWRGHTIQ 360
Qy 361 SRPIGGGLNTSTHGSTNTSINPRLSPFSDRVYWTESYAGVLLWGIYLEPIHGVPTVRFN 420
Db 361 SRPIGGGLNTSTHGSTNTSINPRLSPFSDRVYWTESYAGVLLWGIYLEPIHGVPTVRFN 420
Qy 421 FRNPQNTFERGTANYQSPYESPGLQKDSLTPETTERPNYESYSHRLSHGLISQSR 480
Db 421 FRNPQNTFERGTANYQSPYESPGLQKDSLTPETTERPNYESYSHRLSHGLISQSR 480
Qy 481 VHPVYVSWTHRSADRNTTSSDSITQIPLVKSFNLSGTSVSGPGFTGGDIIRTNVNGS 540
Db 481 VHPVYVSWTHRSADRNTTSSDSITQIPLVKSFNLSGTSVSGPGFTGGDIIRTNVNGS 540
Qy 541 VLSMGLNFNTSLQRYRVRVYAASTQMLRVTVGGSTTFDQGPPTMSANESLTSQSPR 600
Db 541 VLSMGLNFNTSLQRYRVRVYAASTQMLRVTVGGSTTFDQGPPTMSANESLTSQSPR 600
Qy 601 FAEPPVGISASGQTAGISISNNAGROTFFDKIEFIPITATPEAEYDLERAEVNAALF 660
Db 601 FAEPPVGISASGQTAGISISNNAGROTFFDKIEFIPITATPEAEYDLERAEVNAALF 660
Qy 661 TTNTPRLKTDVTDYHIDQVSNLVACLSDFCLDKRELLKVKYAKRLSDERNLQDPN 720
Db 661 TTNTPRLKTDVTDYHIDQVSNLVACLSDFCLDKRELLKVKYAKRLSDERNLQDPN 720
Qy 721 FTSINKOPDFTSTNEQSNFTSIHQSHGWWGSENITIQECNDVFNENYVTLPGTFNECY 780
Db 721 FTSINKOPDFTSTNEQSNFTSIHQSHGWWGSENITIQECNDVFNENYVTLPGTFNECY 780
Qy 781 PTLYQKIGESLKAETRYQIRGVIYEDSDLEIYLIRYNAKHETLDVPGTSEVWPLSVES 840
Db 781 PTLYQKIGESLKAETRYQIRGVIYEDSDLEIYLIRYNAKHETLDVPGTSEVWPLSVES 840
Qy 841 PIGRCGEPNRCAPHFENPDLDCSDRGEKCAHSHHFSLDIDIGCTDLHENLGVWVVK 900
Db 841 PIGRCGEPNRCAPHFENPDLDCSDRGEKCAHSHHFSLDIDIGCTDLHENLGVWVVK 900
Qy 901 IKTOEGHARLGNLEFIEBKPLLGALSRVRAEKWDRKREKLQLETKRVYVYTRAEAVDA 960
Db 901 IKTOEGHARLGNLEFIEBKPLLGALSRVRAEKWDRKREKLQLETKRVYVYTRAEAVDA 960
Qy 961 LFVDSQYRNLQADNTNIGMHAADKLVRIRIYAVLSLSVPGVNAEIFELEGRITTAIS 1020
Db 961 LFVDSQYRNLQADNTNIGMHAADKLVRIRIYAVLSLSVPGVNAEIFELEGRITTAIS 1020
Qy 1021 LYDARNVVKNGDFNGLACNVKGVHDVQOQSHRSVLVIVPEWAEVSOAVRVCPCRGYIL 1080
Db 1021 LYDARNVVKNGDFNGLACNVKGVHDVQOQSHRSVLVIVPEWAEVSOAVRVCPCRGYIL 1080
Qy 1081 RVTAYKEGYGEGCVTIHEIENNTDELKFKNCEEEVYPTDGTGTCNDYTAHQGTAVCNRSN 1140
Db 1081 RVTAYKEGYGEGCVTIHEIENNTDELKFKNCEEEVYPTDGTGTCNDYTAHQGTAVCNRSN 1140
Qy 1141 AGYEDAYEVDTTASVNYKPYEBEETVTDVRRDNHCEYDRGVVNYPPPLPAGYMTKELEYFP 1200

```

```

Db 1141 AGYEDAYEVDTTASVNYKPYEBEETVTDVRRDNHCEYDRGVVNYPPPLPAGYMTKELEYFP 1200
Qy 1201 ETDKWIIEIGETEGKFIIVDSVELLMEE 1228
Db 1201 ETDKWIIEIGETEGKFIIVDSVELLMEE 1228
RESULT 2
US-10-428-961-38
; Sequence 38, Application US/10428961
; Publication No. US20030237111A1
; GENERAL INFORMATION:
; APPLICANT: Baum, James A.
; APPLICANT: Chu, Chih-Rei
; APPLICANT: Donovan, William P.
; APPLICANT: Gilmer, Amy J.
; APPLICANT: Rupa, Mark J.
; TITLE OF INVENTION: Lepidopteran-Active Bacillus thuringiensis Delta-Endotoxin
; FILE REFERENCE: Polynucleotides, Compositions, and Methods of Use (Amended)
; CURRENT FILING DATE: 2003-05-02
; PRIOR FILING DATE: 2000-09-13
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 38
; LENGTH: 1228
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-428-961-38
Query Match 99.8%; Score 6464; DB 15; Length 1228;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1224; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 1 LTSNRKNEININALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVASTVQTGI 60
Db 1 LTSNRKNEININALSIPAVSNHSTQMDLSPDARIEDSLCIAEGNNINPLVASTVQTGI 60
Qy 61 NIAGRIILGVLPAGFAGFASFLVGLWPRGRDQWEI FLEHVEQLINQOITENARNTA 120
Db 61 NIAGRIILGVLPAGFAGFASFLVGLWPRGRDQWEI FLEHVEQLINQOITENARNTA 120
Qy 121 LARLOGLGDSFRAYQOQSLDLENRDDRTRSVLYTOYIALELDFNAMPLFAIRNOEVP 180
Db 121 LARLOGLGDSFRAYQOQSLDLENRDDRTRSVLYTOYIALELDFNAMPLFAIRNOEVP 180
Qy 181 LLMVYQAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVEQTRDYSYCVWEYNTGLN 240
Db 181 LLMVYQAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVEQTRDYSYCVWEYNTGLN 240
Qy 241 SLRGTAASWVRVNOQFRDRLTLGVLDLVALFPSTYTRTPINTSAQLTRREYVYDAIGATG 300
Db 241 SLRGTAASWVRVNOQFRDRLTLGVLDLVALFPSTYTRTPINTSAQLTRREYVYDAIGATG 300
Qy 301 VNMAWMNNYNNAPSFSAIETAVIRSPHLLDLEQLTIFSTSSRWSSATRMHTYWRGHTIQ 360
Db 301 VNMAWMNNYNNAPSFSAIETAVIRSPHLLDLEQLTIFSTSSRWSSATRMHTYWRGHTIQ 360
Qy 361 SRPIGGGLNTSTHGSTNTSINPRLSPFSDRVYWTESYAGVLLWGIYLEPIHGVPTVRFN 420
Db 361 SRPIGGGLNTSTHGSTNTSINPRLSPFSDRVYWTESYAGVLLWGIYLEPIHGVPTVRFN 420
Qy 421 FRNPQNTFERGTANYQSPYESPGLQKDSLTPETTERPNYESYSHRLSHGLISQSR 480
Db 421 FRNPQNTFERGTANYQSPYESPGLQKDSLTPETTERPNYESYSHRLSHGLISQSR 480
Qy 481 VHPVYVSWTHRSADRNTTSSDSITQIPLVKSFNLSGTSVSGPGFTGGDIIRTNVNGS 540
Db 481 VHPVYVSWTHRSADRNTTSSDSITQIPLVKSFNLSGTSVSGPGFTGGDIIRTNVNGS 540

```

seq id no 63 in the provided

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2005, 04:10:31 ; Search time 1515 Seconds
(without alignments)
269.747 Million cell updates/sec

Title: US-10-614-524-2

Perfect score: 6479

Sequence: 1 LTSNRKNEIINALSIPAV.....IGTEGKFIVDSVELLMEE 1228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1424015 seqs, 332791073 residues

Total number of hits satisfying chosen parameters: 1424015

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubppa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubppa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubppa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubppa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubppa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubppa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6479	100.0	1228	15	US-10-614-524-2
2	6464	99.8	1228	15	US-10-428-961-38
3	5912.5	91.3	1227	15	US-10-428-961-63
4	5745	88.7	1228	17	US-10-926-819-8
5	5742	88.6	1228	16	US-10-809-953-10
6	5659.5	87.4	1207	10	US-09-988-462-7
7	5108	78.8	1186	9	US-09-826-660-23
8	3502.5	54.1	1189	10	US-09-972-175-59
9	3502.5	54.1	1189	14	US-10-200-522-59
10	3500.5	54.0	1189	10	US-09-972-175-2
11	3500.5	54.0	1189	14	US-10-200-522-2
12	3496.5	54.0	1189	10	US-09-972-175-61
13	3496.5	54.0	1189	14	US-10-200-522-61

14	3495.5	54.0	1189	16	US-10-782-020-7	Sequence 7, Appli
15	3495.5	54.0	1189	17	US-10-926-819-9	Sequence 9, Appli
16	3494.5	53.9	1189	10	US-09-972-175-4	Sequence 4, Appli
17	3494.5	53.9	1189	10	US-09-972-175-6	Sequence 6, Appli
18	3494.5	53.9	1189	14	US-10-200-522-4	Sequence 4, Appli
19	3494.5	53.9	1189	14	US-10-200-522-6	Sequence 6, Appli
20	3493.5	53.9	1189	10	US-09-972-175-12	Sequence 12, Appli
21	3493.5	53.9	1189	14	US-10-200-522-12	Sequence 12, Appli
22	3490.5	53.9	1189	10	US-09-972-175-8	Sequence 8, Appli
23	3490.5	53.9	1189	14	US-10-200-522-8	Sequence 8, Appli
24	3487.5	53.8	1189	10	US-09-972-175-10	Sequence 10, Appli
25	3487.5	53.8	1189	14	US-10-200-522-10	Sequence 10, Appli
26	3482.5	53.8	1189	16	US-10-809-953-8	Sequence 8, Appli
27	3479.5	53.7	1189	14	US-10-102-469-20	Sequence 20, Appli
28	3476.5	53.7	1189	11	US-09-837-961-6	Sequence 6, Appli
29	3476.5	53.7	1189	16	US-10-825-751-6	Sequence 6, Appli
30	3463	53.4	1181	10	US-09-988-462-11	Sequence 11, Appli
31	3463	53.4	1181	10	US-09-988-462-13	Sequence 13, Appli
32	3463	53.4	1181	10	US-09-988-462-17	Sequence 17, Appli
33	3463	53.4	1181	10	US-09-988-462-28	Sequence 28, Appli
34	3463	53.4	1181	15	US-10-136-998A-4	Sequence 4, Appli
35	3463	53.4	1181	15	US-10-136-998A-8	Sequence 8, Appli
36	3463	53.4	1181	15	US-10-136-998A-10	Sequence 10, Appli
37	3463	53.4	1181	15	US-10-136-998A-12	Sequence 12, Appli
38	3460	53.4	1193	9	US-09-873-873-30	Sequence 30, Appli
39	3460	53.4	1193	10	US-09-916-956A-30	Sequence 30, Appli
40	3460	53.4	1193	10	US-09-997-914-30	Sequence 30, Appli
41	3460	53.4	1193	14	US-10-365-645-30	Sequence 30, Appli
42	3460	53.4	1193	15	US-10-672-163-30	Sequence 30, Appli
43	3460	53.4	1193	16	US-10-739-482-30	Sequence 30, Appli
44	3460	53.4	1193	16	US-10-817-182-30	Sequence 30, Appli
45	3459	53.4	1177	9	US-09-873-873-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1

US-10-614-524-2
; Sequence 2, Application US/10614524
; Publication No. US20040016020A1
; GENERAL INFORMATION:
; APPLICANT: Arnaut, Greta
; APPLICANT: Boets, Annemie
; APPLICANT: Damme, Nicole
; APPLICANT: Mathieu, Eva
; APPLICANT: Vanneste, Stijn
; APPLICANT: Van Rie, Jeroen
; TITLE OF INVENTION: Insecticidal proteins from Bacillus thuringiensis.
; FILE REFERENCE: NEWBTSUS2
; CURRENT APPLICATION NUMBER: US/10/614,524
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: US/09/739,243
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/173387
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1228
; TYPE: PRT
; ORGANISM: Bacillus thuringiensis
US-10-614-524-2

Query Match 100.0%; Score 6479; DB 15; Length 1228;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LTSNRKNEIINALSIPAVSNHSTOMDLSFDRIEDSLCIAEGNNINPLVASTVQTGI 60
Db 1 LTSNRKNEIINALSIPAVSNHSTOMDLSFDRIEDSLCIAEGNNINPLVASTVQTGI 60
Qy 61 NIAGRILGLVPPAGQIAFYSLVGLWPRGRDQWEIFLEHVEQLINQOITENARNTA 120